DOCUMENT RESUME

ED 145 852 IR 005 309

AUTHOR Cameron, William J.
TITLE The Bibliographical Control of Early Books.
PUB DATE Dec 76
NOTE 74p.; Paper presented at the Indian Statistical Institute, Documentation, Research and Training Centre (Bangalore, December 1976); The Sarada Ranganathan Lectures (1976)

EDRS PRICE MF-$0.83 Plus Postage. HC Not Available from EDRS.
DESCRIPTORS Books; *Cataloging; *Data Bases; *Humanities; Information Retrieval; Librarians; Libraries
IDENTIFIERS *Early Books

ABSTRACT

Examples are given of the kinds of machine-readable data bases that should be developed in order to extend attempts at universal bibliographical control into neglected areas, the results of which can be used by researchers in the humanities, specifically those using books printed before 1801. The principles of bibliographical description, organization, and retrieval necessary for the design of such data bases and their control systems are explored and clarified. Theoretical considerations are based upon conceptual development from Dr. Ranganathan's writings and upon the author's experience with the Hand Printed Books Project. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED145852

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE BIBLIOGRAPHICAL CONTROL OF EARLY BOOKS: THE SARADA RANGANATHAN LECTURES 1976

William J. Cameron

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

William J. Cameron
A. Neelameghan

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

IR005 309

Lecture 1

SUBSTITUTES FOR UNIVERSAL BIBLIOGRAPHY

Universal Bibliographical Control is a topic of absorbing interest to library and information scientists who are active in their profession or discipline at the international level. The International Federation of Library Associations (IFLA) devoted its 1972 annual meeting to the topic, and its Office for Universal Bibliographical Control with its emphasis on national bibliographies is one of its strongest, continuing activities. Long before the Federation Internationale de Documentation (FID) was established in 1895, its promoters and progenitors, Otlet and La Fontaine, were actively involved in the field, and their interests have been shared at every meeting. The United Nations Educational Scientific and Cultural Organization (UNESCO) has been strongly promoting schemes for universal bibliographical control such as those conveniently subsumed under the umbrella of UNISIST, or the currently-acclaimed NATIS, the proponents of which, like IFLA, place particular emphasis on national bibliographies and the exchange of bibliographical data in those bibliographies.

Despite all this activity, a search of the literature that it has produced demonstrates fairly clearly that universal bibliographical control as a concept concentrates more especially on current publication, and within that portion of the bibliographical universe (apart from comprehensive national bibliographies) largely on scientific and technical publication. I propose in this series of lectures to broaden the scope of universal bibliographical control by emphasizing the importance of the comprehensive (largely non-scientific) literature of the past. In drawing attention to this vast body of retrospective literature of primary interest to scholars in the humanities, I hope, paradoxically, to re-emphasize the importance of "universal". By studiously ignoring the major efforts now being undertaken in the field, I shall be paying special attention to what may seem to many of my colleagues to be obscure and peripheral, rather than obvious and central. But I am convinced that until <u>all</u> items in the bibliographical universe that have physically survived are taken into account, we shall be dealing with a culturally biased sample, a bibliographical sample that is merely one of a number of possible substitutes for universal bibliography.

In other words, I wish to concentrate upon an area of the bibliographical universe that has been neglected in the major international efforts to develop a system of universal bibliographical control, an area for which existing bibliographical instruments do not yet constitute a viable substitute for a universal bibliography. Nor could it be said that those instruments complement or supplement current international programs for universal bibliography.

Substitutes for Universal Bibliography

I would like to take as my point of departure a quotation from Dr. Ranganathan:

Can we replace universal bibliography by useable substitutes?
Yes. We can.

S. R. Ranganathan, Libri 1953:2:293
"Universal Bibliography and its Substitutes"

I have plucked this characteristically succinct and spirited quotation of Dr. Ranganathan from an article of his where he pitted Subject Bibliographies against National Bibliographies as viable substitutes for universal bibliographies.

In this series of lectures I shall be going even further, and will assert that national and subject bibliographies are necessary, but insufficient, as substitutes for universal bibliography. Dr. Ranganathan wanted us "to choose between what is more helpful but unstable and what is stable but less helpful" and to distribute final responsibility for international Subject Bibliographies evenly among all nations by mutual agreement in order to avoid the instability brought about by war or conflict between nations. He called upon the Federation for International Documentation (and a hoped for merger with the International Federation of Library Associations) to undertake responsibility for the distribution of such responsibility and for establishing consistency in the technique of documentation and for establishing uniformity of standards. And he also called upon such a merged body to "establish a more helpful individualizing and self-perpetuating scheme of classification than the U.D.C." Dr. Ranganathan passionately believed that subject bibliographies could be built consistently (though done by many nations) if international thought could be turned towards the classificatory foundation of documentation. Nevertheless, we can be sure that he would have been willing to entertain ideas about other attempts based on different assumptions that would provide "useable substitutes" for universal bibliography.

I am sure that he would have been very pleased to see the development of INIS, AGRIS, DEVSIS, etc., but I wonder if, like many of us, he would have been pessimistic about the possibility of such international subject bibliographies being extended beyond a few vital areas of international interest. The political persuasion and dedicated professional effort that lies behind the establishment of INIS and AGRIS, resulting in commitment of governments to national support of these information systems in the areas of nuclear physics and agriculture, is greatly to be commended, and if DEVSIS can be accepted by governments on the model of the two established systems, then we may hope that other subject areas may eventually be added. The technological means of ensuring that cooperating nations may contribute to and benefit from the systems -- the exchange of bibliographical data in machine-readable form -- has in fact made feasible what was once only a pipe-dream of proponents of universal bibliography. The computer, with its vast storage capacity, and its ability to manipulate bibliographical data so as to provide a multiplicity of access points to it, has changed the technological environment of documentation even if the fundamental problems remain to be solved. As a result of the new possibilities created by electronic data processing, one cannot quite so confidently agree with Dr. Ranganathan when he states:

> The world of print had left the region of five digits even in the sixteenth century. M.V. Iwinstey had estimated that by the turn of the present century the number of books in the world had gone well into the region of eight digits. Indeed the annual output of the world itself has entered six digits. This alone is enough proof that a universal bibliography is impracticable.

Nor need we agree with earlier categorical statements such as that of Frederic Alphonsus Ebert that "external impediments" are "insuperable".

> ...it must be acknowledged to have been a weakness of mind, even disregarding the insuperable external impediments, that some literary persons dreamed of a general bibliography, in which no compendium, no pamphlet or collection of verses should be passed over.
>
> A General Bibliographical Dictionary, 1838, p.iii

Nevertheless, useable substitutes for universal bibliography are preferable on grounds other than unwieldiness, for modern techniques of data storage can even cope with the monstrous bibliography which Dr. Ranganathan describes thus:

> A bibliography to be universal should include all published materials, whether books or parts of them or periodicals or serials or articles in them or any combination of them, on all subjects, in all languages, in all countries and at all times. Of these four factors, time gradually swells the number of items to be included.

If we were to apply the principle of "Every book its user", a highly-sophisticated retrieval system or systems will be needed to deal with such a monstrous bibliography.

Despite mind-boggling warnings about size of, and accessibility to, such a universal bibliography, I believe that useable substitutes can be devised that would in fact be contributions towards a useable universal bibliography. First necessity would be a steady accumulation of "stable but less helpful" bibliographical entries in computer storage, and second necessity would be a proliferation (under computer controls) of "helpful but unstable" peripheral devices for providing specialized bibliographies of many different kinds.

I like to think of the completed universal bibliography as a large and complex jigsaw puzzle in which two kinds of pattern are discernible. First is the regular but useless pattern of the pieces themselves -- a pattern discernible from the reverse or uncolored side of the puzzle. The other kind is the meaningful patterns revealed on the picture side of the puzzle even when only parts of the puzzle have been completed.

Our aim is to permit the viewer of the completed puzzle or even of completed portions of the puzzle to enjoy those parts of the second kind of pattern that are most to his taste. But we will have to put together the pieces according to the way they contribute to the first kind of pattern.

The first step in creating this huge jigsaw puzzle is to decide which areas of the puzzle are best (or easiest) to fill in; the second is to identify those pieces that can be best made to interlock to form those areas (the useable substitutes); and the third is to make sure that the colorful patterns revealed when the pieces are put together will be brought to the attention of those who can best use them. Those patterns, which are the main reward for the patient jigsaw puzzle fan, would be, in our simile, useable bibliographies different from the useable substitutes for universal bibliography that went to make up each piece of the puzzle.

In Documentation and its Facets, Dr. Ranganathan spells out five characteristics which might form the basis for restricting an "omnibus universal documentation list" to useable lists.

1. Area of origin of the documents
2. Subject of the documents
3. Period of production of the documents
4. Standard of the documents
5. Service aimed at by the documentation list.

To these I would like to add a sixth:

6. The physical process by means of which the documents were produced.

When one considers these 6 "characteristics", one notes that two principles might be at odds with each other; 1, 3 and 6 (i.e. area of origin, production period, and physical means of production) imply decisions of a predominantly objective, non-teleological kind. Two, 4 and 5 (i.e. subject, quality, and user needs) imply a selectivity based on knowledge in the minds of decision-makers about the purposes for which the lists will be used. I believe that useable lists must first of all be limited according to area, period, and means of production, and stored in machine-readable form according to conventions that will enable selective mechanisms to be developed in order to make such lists truly useable by people with a wide variety of purposes. To use my jigsaw puzzle metaphor, the pieces must be shaped to interlock before the patterns begin to emerge. The lists must be comprehensive within the limitations set, so that the greatest "bibliographical control" can be experienced through the selective mechanisms. An understanding of the notion of "bibliographical control", a useful, but somewhat undefined (and perhaps indefinable) term, will place the notions of universal bibliography (or its "useable substitutes") and the notion of selective retrieval into a new focus, and perhaps enable me to refine on Dr. Ranganathan's solutions to the problem of achieving universal bibliographical control.

Bibliographical Control

A very intelligent philosphical view of the concept is provided in Patrick Wilson's Two Kinds of Power,[1] which is a Carnapian attempt at explication of the notion of "bibliographical control". A more pragmatic, historical, and sociological view of the framework is provided by Margaret E. Egan and Jesse H. Shera in Prologomena to Bibliographic Control.[2]

Wilson attempts[3] to formulate a concept (or explication) of "bibliographical control" which will be similar to the explicandum (i.e., to the common use or uses of the term), but which will also be "exact" and "fruitful" (where "fruitful" means "useful for the formulation of many universal statements, i.e., empirical laws or logical theorems"). He also wants this concept to be "as simple as possible". After defining the bibliographical universe over which bibliographical control is or might be exercised as "writings and recorded sayings" (cf. Ranganathan's "all published materials"), he analyzes the individual items in that universe into a work, a text, and an exemplar of the test. An exemplar (which term Wilson prefers to "token" or "replica") may be either a performance or a copy of a text, where "copy" is extended to include a text stored up in the memory of a man or a machine. (He nevertheless excludes "pieces of music and pictures" from the bibliographical universe). He then goes on to deal with the notion of a unit as an item of information rather than a text. He rejects somewhat hurriedly the opinion that this notion cannot be "a generally advantageous procedure, however useful it might be in particular and limited circumstances".

In this respect, the philosopher may have overlooked the theoretical as well as practical advantages of considering the concept of "accessibility" which is explained in historical terms by Egan and Shera. They emphasize the essential difference between content accessibility and physical accessibility to information, in such a way that Wilson's ill-considered rejection of the notion of an "item of information" seems inadvisable. One might note, in passing, that Ranganathan's insistence on attention to the subject classificatory foundation of documentation is very closely bound up with the notion of content accessibility. He claims that "nearly 40 percent of time and thought-energy in documentation work" could be saved by a useful classification scheme. However, a more philosophical explication of the concept of "accessibility" as put forward by Egan and Shera seems warranted on reading their conclusions about "new types of classification" introduced by "the development of new mechanical devices for quickly sorting large numbers of coded references without regard to the original arrangements":

> The entire problem of classification, then, presents

---

1. Patrick Wilson, Two Kinds of Power, An Essay on Bibliographical Control (Berkeley and Los Angeles, Univ. of Calif. Press, 1968).

2. Journal of Cataloging and Classification V, 1949, 17-19.

3. Wilson, op. cit. p. 24.

three aspects:

a. Classification of physical units of recorded information within a single collection, whatever physical form such units may take.

b. Classification of subject content, based upon adequate subject analysis and divorced from the physical unit either of book or of single library collection.

c. Identification and description by code of small units of information, regardless of relationships.

These three "aspects" are puzzling, even confusing. Why should "physical units of recorded information" be limited to a single collection? Is it possible to divorce subject content from a "physical unit"? (Or do the authors have in mind the classification of knowledge in the mind of the information seeker?) Why should relationships between "Units of information" be disregarded? Egan and Shera seem to separate notions that are closely connected, and seem to combine notions that are better or more fruitfully kept separate.

The concept of "physical accessibility" probably stems from the basic power exercised by an information-seeker when locating and actually obtaining an exemplar of a text or texts that he needs; the concept of "content accessibility" stems from the power he exercises when identifying where in those texts (or, more specifically, where in a particular exemplar of a particular text of a work) the information he needs is to be found. Neither power, however, can be exercised without (a) knowledge either in the user or his agent of the universe of subjects; (b) a physical storehouse (library, archive, tape library, etc.) and (c) a (physical and intellectual) inventory or catalog containing both surrogates for the physical units and surrogates for the subject knowledge of the user or his agent. The physical surrogates can be called "bibliographical entries"; the surrogates for subject or content knowledge are explicit or implicit "content" classification schemes, which may be (and usually are) modified in certain ways by physical classification. (e.g., the physical size of a book may determine its physical location, and therefore, its physical relationship with other books). "Physical access" therefore implies locating both physically and mentally, the actual vehicle of information; whether that vehicle is a single physical unit or several units, physically together or separate. "Content access" implies identification (and sometimes location) of information carried by the physical vehicle. No doubt a philosopher could further clarify the distinctions here made.

I am convinced that all devices for exercise of bibliographical control must be based on the assumption that "physical access" is of primary concern, for only when the "physical unit" is defined and located, can "content access" be meaningfully provided. The physical unit may be, but is not necessarily, the carrier of a single text (Wilson's unit for the bibliographical universe). Subject analysis may depend on further analysis of the physical unit, or even the bringing together ("collocation") of related physical units. To put it another way, the "identification and

description by code of small units of information" (which might be interpreted by some as classification, of which indexing is a part, and by others as indexing, of which classification is a part) can only be done by reference to location of the small units either in a single physical unit or a collocated group of units. I would like to illustrate by reference to my own experimental system for bibliographical control of early books -- the HPB project -- which I will describe in more detail later in this series of lectures.[1]

The "Bibliographical Universe" and its Units

The HPB project, in Wilson's terms, has as its bibliographical domain "all writings and recorded sayings" for which exemplars made before 1801 are still extant (i.e. it is limited by period of production but not by area). For practical reasons (as well as Wilson's theoretical reasons) all pieces of music and pictures are excluded, and for purely practical reasons, all exemplars not made by means of letterpress impression are excluded (i.e. it is limited by means of production). Thus, books consisting entirely of engraved sheets are excluded. For obvious reasons, texts recorded in the mind of human beings before 1801 but not published until after 1800 are excluded, and (at least for the foreseeable future), all manuscripts. Periodical publications present a particular practical problem not covered by Wilson's theoretical considerations, but are feelingly identified by Egan and Shera as historically the cause of "the first crisis in bibliographic control". Dr. Ranganathan put it in these words:

> An additional factor driving universal bibliography beyond the region of practicability took

---

1. Fuller descriptions of the developing project will be found in:

William J. Cameron and Brian J. McMullin, The HPB Project: Phase I. Describing an experiment in creating A Universal Bibliography of Hand-Printed Books by using the computer's memory bank and by developing a step by step cumulation of existing records. London, Ontario, School of Library and Information Science, University of Western Ontario, 1968, 20, (158) pp.

William J. Cameron, Brian J. McMullin and Joginder K. Sood, The HPB Project: Phase II. Describing an experiment in creating a computerized cumulative short-title catalog of Hand-Printed Books leading to a universal bibliography of books printed 1453-1800. With illustrative material from John Milton and Miltoniana, London, Ontario, School of Library and Information Science, University of Western Ontario, 1970, ii, 165 pp.

William J. Cameron, Peter Gulutzan and Tom Delsey, The HPB Project: Phase III. An exploration of ways of improving "The National Union Catalog: Pre-1956 Imprints" by use of the machine-readable records of the HPB project, standard bibliographies, and published library catalogs, London, Ontario, School of Library and Information Science, University of Western Ontario, 1976, IV, 183 (77) pp.

> shape early in the seventeenth century when a new species of publications now called periodicals was started. Analytical entries of individual articles have always been assumed as part of universal bibliography.

In his terms an "individual article" forms what Wilson would call a text (or its exemplar).

The difficulty of defining in practice what should constitute a single bibliographical unit for the HPB project led to the development of the notion of "a bibliographically distinct volume". Two manifestly imprecise definitions were put forward[1]: (a) a book pamphlet, or other printed item conceived of and printed as a single physical entity; (b) a volume consisting of a title page (or its equivalent) followed by a complete series of signatures and pagination. The first definition invites the decision-maker to entertain a notion presumably held by the producers of the volume, the second to entertain the notion as deduced by an examiner of the physical product.

It will be noted that a distinction is thus being drawn betwern a "physical unit" of the bibliographical universe and a "bibliographical unit" of that universe. In practical terms, the distinction is important for items other than periodicals. An example might be taken from Milton. A work by him with the title A complete collection...[2] is usually found in the form of three bound volumes (physically separate) as described on the title page of the first volume. The third volume is intellectually separable by its language (it is in Latin, whereas the first two volumes are in English), but its contents are of the same kind as the first two volumes (i.e., a variety of rhetorically distinct texts covering a variety of subjects in the form of prose). But bibliographically and to a certain extent, intellectually, the third volume is merely an appendix to the first two volumes, which are themselves bibliographically only one volume because there is a single series of page numbers and a single series of signatures. (The third volume, or appendix, has a series of signatures prefaced with an

---

1. The HPB Project: Phase II, p. 2.

2. The British Museum catalog describes it thus:

> A complete Collection of the historical, political and miscellaneous works of John Milton, both English and Latin. With some papers never before publish'd. To which is prefix'd the life of the author (by J. Toland). 3 vol. Amsterdam, (London) 1698, fol. 713. K.8.
>
> The pagination of these volumes is in two series: one of the English, the other of the Latin works. There are several distinct title pages, variously dated from 1694 to 1698.

asterisk, indicating that they are the second series within a complete series.) From the bibliograpnical point of view, then, the three physical volumes are one volume; from the rhetorical or subject point of view, the text of this single volume consists of 35 texts grouped by means of seven distinct title pages,[1] which cover four subjects and three miscellaneous categories. Content access to each text is solely through the title or chapter headings, as no index or other device of content analysis is provided except for a listing of 35 numbered items "in true order" of composition which is provided on pp. iii-iv of the so-called Volume I.

The distinctions made in the paragraph above, with the implied definitions of the underlined words, should help us to cope with the problem of physical accessibility (and thus, eventually, with content accessibility).

Just as content accessibility to the 35 texts in the Milton volume is made possible by identifying the bibliographical unit which carries them, so by identifying the bibliographical unit, content accessibility is made possible to the 555 texts (or "rhetorically distinct units") which are contained in any edition of Steele and Addison's Spectator essays. Typically, the dozens of editions of the periodical essays that issued from the press after 1712 consisted of eight physical volumes which would be bought as a single work. But the essays were originally issued separately, each on a single sheet printed on both sides with the following typical heading on one side:

NUMB. 11

The SPECTATOR

-Ast Alii sex
Et plures uno conclamant ore - Juv

Friday, March 2, 1711

On the verso, beneath the text (printed in two columns) would appear an imprint similar to the following (Jacob Tonson's imprint appeared in fairly

---

1. They occur on signatures Bi, Tl, Nnl, Llll, Ppppl, 5Ml, and the unsigned leaf before *Al.

regular alternation throughout the printing history of the 555 sheets):

> LONDON: Printed for Sam. Buckley, at the Dolphin in Little-Britain; and Sold by A. Baldwin in Warwick-Lane, where Advertisements are taken in.
>
> R

Now, according to the concept of a "bibliographically distinct volume" a numbered series of physically separate items (such as Steele and Addison's Spectator) without page numbers or signatures is not strictly a completed volume until it is reprinted as a single physical entity with a title page, signatures, and, possibly, pagination to provide evidence that it was conceived of as a "bibliographically distinct volume". Thus the original folio sheets need not be included in our "bibliographical universe" because they are bibliographically only part of a whole even though they are physically distinct. In the aggregate issue,[1] however, a title-page would be sufficient evidence of completion of the intended volume, and (like the physically reprinted forms which continued to appear throughout the 18th century) should be considered a "bibliographical unit" within our universe.

However, completeness of text (whether or not that text is issued serially in separate physical entities) is not the major criterion for a bibliographical unit. Individual folio sheets (either broadsides or broadsheets) of complete texts must be considered to be "bibliographically distinct volumes" only if there is no indication (either in that physically separate entity or in a subsequent one) that they are part of a series; and a physically separate volume of a multi-volume text (whose completeness is indicated by continuity of theme from one volume to another without the intervention of another text) may not be a bibliographically distinct volume. For instance, each volume may have a new series of signatures and pagination, or a four-volume set may be made up from two physically distinct volumes with a single series of signatures and pagination followed by two volumes each with a separate series of pagination and signatures.

Further explanation need not be undertaken here of the implications of establishing a "bibliographically distinct volume" as the unit of our limited portion of the bibliographical universe rather than the textual one of "rhetorically distinct unit" or the mixed content one which seems to underly most cataloging codes. But the exclusion of parts of a series and of items not produced by letterpress printing which the concept entails has considerable bearing on how the units within our limited portion of the universe are to be described.

Traditional cataloging practices are different for monographs, periodicals, and manuscripts. All we need be concerned with is adaptation based on sound theoretical principles of practices used in describing monographs.

---

1. For a detailed explanation of this bibliographical term, see W.J. Cameron "A Late Seventeenth Century Scriptorium" Renaissance and Modern Studies (University of Nottingham) VII, 1963, pp. 25-52, especially pp. 27-30.

The principles to be derived for a universal bibliography to be stored in a computer must include economy of data in the entry. The greatly limited value of minimal amounts of data for each unit must be compensated for in the peripheral control devices by maximizing the possibility of adding appropriate data where needed. Two levels of control are in fact implicit. The macroscopic value of the universal bibliography must be supplemented by microscopic features in the control devices. The process of production (letterpress printing) makes it possible to subsume many thousands of exemplars under one brief entry. The process of producing manuscripts makes it difficult to subsume exemplars under anything except similar textual content, and the uniqueness of most manuscripts requires an almost microscopic level of description. But I shall not further complicate the distinction with examples, for such explanations require the physical presence of some of the books to be described. Our unit of the bibliographical universe has proved to be fairly easy to identify in the past few years, when cataloging some 50,000 or so books.

Manuscripts in fact present a similar problem to the one identified by Egan and Shera as the cause of the second phase (= crisis?) in bibliographic control: "the increasing publication by government agencies or corporations of an extensive literature whose content was of scholarly importance but was frequently accessible only through the listings and classifications of the issuing office". The usual term used for such items of corporate authorship in the bibliographical universe is "documentation" because such a large proportion of them are not in conventional (letterpress or rolling press) printed form. Because Egan and Shera equate their concept of "bibliographical control" with the European use of the term "documentation", we have some evidence that the new concept was born of this newly-recognized practical problem. The problem has always existed, however, but the users of archival materials have never felt the same urgency and desperation as the potential users of "documentation", so the problem of content access and physical access to manuscripts has always been of second priority. First priority since the mid-15th century has always been access to letterpress-printed materials. The HPB project will take up the problem of both manuscript and periodical publications only after the first priority has been dealt with. Both MSS and periodicals constitute different pieces or sets of pieces of our jigsaw puzzle.

Access to Described Units of the Limited Universe

The heart of Egan and Shera's attempt at a prolegomena to bibliographical control is their professional awareness of the necessity for producing appropriate tools for practising the profession of librarianship (other than physical curatorship). In keeping with this highly practical approach, they define "bibliographic controls" as "devices" or "tools" by mechanistic analogy. "In mechanics, a control is a device by means of which the energies of a machine may be directed to the achievement of a given end with a minimum of effort." Cybernetic implications are perhaps not fully recognized when they conclude that bibliographic controls are "those devices used to direct intellectual energy in extracting from the totality of recorded information those portions relevant to a particular task with the utmost speed and economy".

On the other hand, the heart of Wilson's attempt at analysis of bibliograpical control is his maintenance of "our common-sense distinction between powers whose exercise requires appraisal and those whose exercise requires only bare description". He maintains the theoretical distinction between "exploitative" and "descriptive" power by arguing that the second is an inferior substitute for the first, while admitting that the limited powers actually possessed by people are complex mixtures of the two kinds. He rebuts three possible objections to the distinction:

1. that the difference between the two becomes negligible (as trivial as would be a division of descriptive control itself into a thousand different sorts of power) if by means of a "thought-experiment" we imagine what it would be like to have either sort of power in the highest degree;

2. that the distinction might disappear if requests framed in terms of suitability of means to ends could always be replaced by (if not "reduced" to) requests framed in terms of some combination of neutral descriptions;

3. that the difference is not a difference of kind, but a difference of degree along a single scale between neutral descriptions and appraisals or evaluations.

In not pursuing the arguments against these objections, Wilson (in an act of "exploitative" bibliographical control?) refers us to the most valuable of modern Anglo-American writings on value-theory and connected subjects (p. 32).

Wilson, however, does not take into account the major objection to his distinction. One might reasonably object that the act of appraisal cannot take place without a prior act of description, and that the power of appraisal is likely to be fundamentally distorted in any situation where accurate description is lacking. How can we appraise Wilson's exploitative bibliographical control of Anglo-American writings on value-theory without a comprehensive description of all or most of those writings? How can we compare Wilson's selection of such writings with the more comprehensive list (and the list must be based on selection, which implies a degree of appraisal) unless we display the two lists within comparable (i.e., neutral) descriptive conventions?

Would not the neutral description of the products of other exercisers of bibliographical control with similar purposes to our own increase our power to exercise our own control? Would not we be able to extend our "exploitative" power by bringing together, in one list, items described within similar conventions from a multitude of lists compiled for different (or even similar) purposes? Would not our opportunity for appraisal be improved if fairly neutral acts of selection were made from unappraised listings of items in our "bibliographical universe"?

With these thoughts in mind, I would like to return to my earlier statements

about universal bibliographical control, and assert that

1. useable substitutes for universal bibliography should consist of neutrally descriptive items within objectively definable limitations,

2. that control devices applied to the evolving universal bibliography should maximize or optimize the exercise of the appraisal powers of potential users or their agents.

This two-part approach to universal bibliographical control should make it possible, with the computer technology available to us, to utilize all the bibliographical work of the past in one of two ways - to add to the cumulative universal bibliography, or to select from it useable portions in suitable arrangements.

Selection, of course, is not enough. Ordering of selected items is also part of bibliographical control, and may exhibit a conventional, "descriptive", or "neutral" quality (such as alphabetical order, chronological order) or be the result of "appraisal" and a sense of specific purpose. The bibliographical tools that a manager of a system of universal bibliographical control must be able to produce must consist of devices not for retrieving the interlocking pieces of the jigsaw puzzle, but for retrieving the properly selected and properly displayed patterns on its surface.

Exploiting the Described Units of the Bibliographical Universe

Now that we have defined the limited part of universe over which we hope that the HPB project's control devices will have power, and commented on the validity of Wilson's "common-sense distinction" between appraisal and description, I would like to respond both theoretically and practically to the key question at the beginning of Wilson's Chapter II (Describing and Exploiting):

> What might a person be able to do or have done to things in the bibliographical universe that would count as exercises of the power we call "bibliographical control"?

In theory "a person" may be anyone; in practice we will find that it is a very limited class of persons who would wish to have power over the limited bibliographical universe of the HPB project (i.e., books produced by letterpress printing befor 1801). This class of person is largely represented by a specialized scholar in the humanities who "thinks he can attain, or get some way towards attaining, a goal by means of reading". The goals will also be found to be very limited and should be described in much the same way as items in the bibliographical universe were described by Wilson. Such a scholar would indeed "like to be able to have the best means available for the attainment of the goal, to have what we can call the best textual means to his end". [1]

---

1. The quotations are from Wilson, p. 21.

We may wish to add to the concept of "the best textual means" a prior consideration - the best means to the best texts, and these are properly treated by Wilson as "bibliographical instruments". Quite often, however, a careful practical distinction should be made between various categories that are included within the theoretical concept of "textual means". Reading is not such a simple means to a man's goal as is assumed in Wilson's treatment of the subject.

"Best textual means" implies selectivity - the selection from the bibliographical universe of a particular item or items most helpful to the scholar in attaining his goal. The power to select an item or items can only be exercised when several items are presented for appraisal; the highest power of selection can be exercised only if all possible items are presented for appraisal. In order to have such power to the highest degree, one must be able to locate or 'assign to a position' all items in the bibliographical universe. So we return to the notion of universal bibliography or its useable substitutes. In order to locate all items, each item must first be described - and the method of description (I would like to assert) is dependent upon (1) the nature of the item to be described (in our case, a representative exemplar of all exemplars of all texts), (2) the use to which the item is to be put (viz. "reading") in attaining a goal (in our case, the goals of a limited class of scholars in the humanities), and (3) the nature of the "bibliographical instruments" by means of which the "best textual means" are to be provided (in our case, the printouts or cathode-ray-tube displays from the electronic machine called a computer).

"Best textual means" also implies relationship - the bringing together of different items from the bibliographical universe that will provide better help to the scholar in attaining his goal when linked than when apart. The power to relate several items is dependent on the power to select those items, so this collocative or relational power may, in one sense, be taken to be a secondary or dependent power. Nevertheless, it is essential to pay attention to the order or pattern in which bibliographic surrogates for texts (or their exemplars) are presented as a result of the exercise of bibliographical control.

The great advantage of separating the function of a universal bibliography or its parts from the products of bibliographical control devices is that the universal bibliography can be stored in any order or pattern convenient for our machines. But the products of the control devices must be arranged (as well as selected) to optimize the use to which the individual reader wishes to put them.

In the next lecture, I propose to outline how these theoretical considerations are being worked out in practice in my own experimental HPB project. I will draw atention to the ways in which the tradition of short-title cataloging has been adapted to provide contributions to a universal bibliography, and explain how the problems of selection and ordering are being tackled by use of control devices. In subsequent lectures, I hope to broach the problem of identifying and categorizing the purposes of scholars by illustrating how some of those purposes can be served by some of those control devices and by the substrata of bibliographical tools that are an essential complement to the macroscopic tools that are the main concern of library and information scientists active at the international level. The implications for national and international planning will, I hope, emerge as I proceed.

Lecture 2

SHORT-TITLE CATALOGS AS "USEABLE SUBSTITUTES" FOR UNIVERSAL BIBLIOGRAPHY

In discussing bibliographical control of early books in my previous lecture, I hinted at the existence of two levels of control which a scholar would like to exercise, two levels which we might call macroscopic and microscopic. The two are often in conflict in the work of a particular bibliographer or library cataloger for quite understandable reasons. The functions of a macroscopic "bibliographical instrument" and a microscopic "bibliographical instrument" cannot realistically be combined, but some coordinating device between them can in fact be developed if the two functions are clearly differentiated. In the next three lectures I shall be dealing with the macroscopic - the comprehensive listing of an exemplar of every text within limited portions of the bibliographical universe, together with associated classificatory or retrieval systems that permit the scholar to find within those listings an item or group of items arranged in a manner relevant to his purposes. In the final lecture I shall be dealing with the microscopic - the comprehensive listing of every surviving exemplar of one text (or a small number of related texts) together with the associated classificatory or retrieval systems that enable a scholar to use the resulting information to help him or her accomplish his purposes. The microscopic must be considered secondary to the macroscopic, as the bibliographical tools listing one exemplar (or more) are antecedent to those tools which list all exemplars (or as many as needed) of a particular text. We might take as an example such a limited field as European incunabula, a field well-covered by bibliographers and catalogers. We shall find such monumental efforts as the Gesamtkatalog der Wiegendrucke, based as it is upon short titles derived from Hain's Repertorium Bibliographicum, making very slow progress, and requiring considerable supplementary work. The first volume in 1925, containing 1256 items, covered Abano-Alexius. The eighth volume, published just before the 1939 war halted the enterprise, contained only 475 items and covered Eike-Fredericus. Even Hain has numerous supplements and commentaries (Coppinger et alia).

Comprehensive Catalogs and Their Useable Substitutes

If we turn to a much broader field, and define our term "early books" as all books printed before 1801, we shall find attempts at meticulously detailed description much rarer than during the incunabula or fifteenth century period. Attempts at comprehensive listing tend also to be less frequent. But perhaps the most comprehensive bibliographical tools ever produced are catalogs of the largest libraries that have been printed for use outside the library; and union catalogs of several libraries, especially those published for use outside the libraries whose holdings are listed in them.

The publication containing the greatest number of items for the pre-1801 period is undoubtedly The National Union Catalog: Pre-1956 Imprints, but no estimate of the total number of such items has yet been made, partly

because only two-thirds of the catalog has as yet been published, and partly because statistical techniques for measuring such heterogenous material leave much to be desired. It is probable that the British Museum Catalogue contains more English imprints than NUC, and the Bibliothèque Nationale's catalog probably contains more French imprints, but statistical analysis of such multi-volume publications is also difficult to undertake, and no comparative estimates of pre-1801 items in them are available. Such elementary quantitative estimates are difficult to make simply because the desired selection of the items that are to be measured is scattered almost arbitrarily among the remainder. The arbitrariness is a result of shifting changes in acquisition patterns over the history of a particular library or group of libraries, and fragmentation a result of alphabetical arrangement by author. It is only when the heterogenous nature of the whole catalog has been homogenized (by a deliberate act of selection) into more specific (and useful) catalogs with more limited coverage that quantitative estimates (a crude form of bibliographical control) can be accurately made. Dr. Ranganathan's notion of "useable substitutes" for even the limited universe of such catalogs thus becomes a form of bibliographical control.

Fortunately for us, some useable substitutes for the General Catalogue of the British Museum have in fact been made. The simplest form of them is the single author catalog - merely reprints from the appropriate sequence in the General Catalogue. But of more value for our present purposes are the series of "short-title catalogs" that have been appearing since 1884.

A short history of so-called "short title" cataloging will explain why the term has come to be accepted, and why English imprints have been more thoroughly submitted to such methods of bibliographical control. In 1884, the Trustees of the British Museum published in 3 volumes a Catalogue of books in the library of the British Museum printed in England, Scotland and Ireland, and of books in English printed abroad to the year 1640. George Bullen, the compiler, had selected the entries from the British Museum's General Catalogue of Printed Books and had cut back the size of each entry partly by abbreviating the title in each - hence the common term "short-title" catalog. Since 1884, a series of similar catalogs for other European nations (or language groups) before 1601 have been published.[1] But more importantly, the obvious usefulness of Bullen's 3-volume compendium to the individual scholar (who could not possess a copy of the huge General Catalogue) and its limitation to the contents of only one library, led the Bibliographical Society (founded 1892) to encourage its members to compile what eventually was published as A short-title catalogue of books, printed

---

1. e.g., for the Romance language countries: Spanish 1921, French 1924, Portuguese and Spanish America 1926, Portuguese 1940, Italian 1958. The Spanish and Portuguese volumes have been carried forward to 1700 by V.F. Goldsmith's STC (1974). The Italian one has been used to compile an STC (ed. Robert G. Marshall) of such books held in selected North American libraries (3 vols, 1970). Similar catalogs were produced half a century ago by The Library of the Hispanic Society of America and other interested cultural institutions, but suffer from lack of reference devices.

in England, Scotland and Ireland and of English books printed abroad 1475-1640, 1926. This one-volume compilation, familiarly called "the STC" or (after its editors and principal compilers) "Pollard and Redgrave" firmly fixed the term "short title catalogue" and its acronym in bibliographic parlance. Revision of this enormously influential book was begun by the late W.A. Jackson and F.S. Ferguson, and has been brought to publication stage (Volume II of the two-volume work was published in the summer of 1976) by Miss K. Pantzer at Harvard University. The final publication will represent the most extensive and comprehensive listing that has ever been produced of surviving items limited to a well-defined period, well-defined area, and a well-defined means of production.

Donald Wing continued the work of short-title cataloging 17th century English books (without the prior publication of an STC of books of the period in one major library) when he published between 1945 and 1951 his Short-title catalog of books printed in England, Scotland, Ireland, Wales and British America and of English books printed in other countries, 1641-1700. A second edition, not elaborated quite so extensively as the revision of Pollard and Redgrave's STC, began to appear in 1972. The reinforcement of acceptance of the term "short-title catalog" is noteworthy, but so also is the extension of the geographical area. When an 18th century STC is produced, geographical coverage will undoubtedly result in a significant increase in size if the tradition is followed. But tradition, though probably a very strong influence, will not be followed in every respect, for literary growth and the freeing of the means of production of printed works brings with it a wide diversity of materials and ephemera that caused the original STC compilers and even Wing little trouble in defining the scope of their portions of the bibliographical universe. On the other hand, the fact that very few STCs for non-English materials have yet developed out of the British Museum's STCs for non-English materials in their collection has perhaps caused advocates of an English 18th century STC to avoid taking a more global approach to the task.

Ever since the publication of Wing's STC, voices have been raised to continue this means of bibliographical control to the end of the 18th century - a date that can be justified on technological as well as cultural grounds, for 1800 represents the beginning of the period when machines begin to take over from hand-operated printing presses, and machine-made paper begins to supplant hand-made paper. It now seems likely that the world of scholarship, librarianship, computer technology, and reprography are about to join forces to fill this desideratum. Prompted by a survey of scholars' needs for bibliographical tools, the American Society for Eighteenth Century Studies, under the leadership of its Executive Secretary, Professor Paul Korshin, began last year to organize both interested members and expert non-member librarians, bibliographers, computer experts, and scholars with the aim of mounting a conference to study the feasibility of producing an 18th century English short-title catalog.

A committee of knowledgeable people in Great Britain was formed, and a similar committee in the United States and Canada. As a result of discussions on both sides of the Atlantic (including a special seminar at the Seventh Annual Meeting of the American Society for Eighteenth Century Studies at the University of Virginia in April 1976), and with funding from the

National Endowment for the Humanities, a very fruitful meeting of 40 distinguished librarians, bibliographers, and other scholars, together with computer cataloging specialists, took place in the British Library from June 14-18, 1976. The results of the week-long deliberations were positive: such a catalog was deemed to be feasible and desirable, and an organizing committee was appointed to explore further a number of issues raised at the conference, in particular the procedures necessary for making a practical start on the catalog itself.

The general principles established at the meeting will probably be modified after the Organizing Committee's investigations are completed, but the consensus of opinion seemed to be that the STC should include all books printed in the 18th century in the English language, and in vernacular languages wherever printed, and all books printed in the British Isles and British North America in whatever language. All categories of printed material which are readily definable and recordable without recourse to exhaustive research was thought to be a good working principle for inclusion of an item in the STC, but one senses that the categories for exclusion (serial publications, maps, and engraved music with no letterpress text present) will be augmented by other categories from "the ragged edge of 18th century publishing" after the Organizing Committee has given the matter further thought.

A test sample will be run in the British Library to try to cut the Gordian knot created by conflicting views about the nature of the STC entry. Rather fuller entries than would ultimately be needed, with eventual abbreviation for output display was advocated by some experts; others advocated the definition of the STC formula at the very outset, thus requiring the recording of only the minimal amount of data for each item. (You will no doubt be able to guess by the end of this lecture which side of the debate accords with my opinion and experience). Some minimal standards were suggested, and compromise solutions to problems based on the peculiarities of 18th century printing and publishing practices were realistically anticipated.

Computer technology was accepted as essential to the enterprise, and the advantages and disadvantages of working with such machines and their peripheral equipment (and reprographic technology, too) would undoubtedly determine some of the cataloging methodology.

Strategies for creating the database seem to have fallen into two categories (not necessarily mutually exclusive). Conversion of the British Library General Catalogue to machine-readable form compatible with MARC format (the Library of Congress' machine-readable cataloging system) as a prior project, with subsequent modification and emendation of the machine-readable records was strongly advocated, and as strongly opposed, especially by those who advocated prior establishment of an STC formula and the gathering of only so much data as was necessary to satisfy those minimum standards. Both British Library and MARC records contain such a large proportion of entries that would need emendation by comparing the entry with a copy of the book described, that the need to examine actual copies could be the single most important determinant of strategy for data gathering. Other strategies, such as conflating marked-up copies of NUC: Pre-1956 Imprints and GK3 (the latest edition of the British Library Catalogue) before either was converted

to machine-readable form were advocated so that the estimated 25% of items that could not easily be matched because of inadequacies in existing catalog records could be dealt with first, and priority also given to making less readily-accessible collections more so (e.g., those not adequately represented in NUC). Some others assumed that those libraries whose 18th-century records were already in machine-readable form either wholly or in part, would be given priority for incorporation into the database. But it may be that insufficient data would be available. It was obvious that most of the group wished to avoid cataloging uncataloged material, but it may very well prove true that this is where much of the material not in NUC or GK3 is to be found.

Manpower for the project was, of course, carefully discussed, but the resolution of problems of strategy would undoubtedly influence the final decisions as to host institution or institutions for the project, size and composition of staff for the project, and their relationship to contributing institutions and staff.

As a participant in this very interesting and productive meeting, I was very pleased that my own opinions and beliefs (based on many years' experience with STC problems) were listened to courteously and open-mindedly, and were seriously debated, then carefully incorporated into the range of possible strategies. The Organizing Committee is undertaking a careful investigation of all possible strategies, and I hope that the best and most economical ones will eventually be adopted. Nevertheless, I am concerned about a global approach to short-title cataloging, and will continue to watch from the sidelines for an international strategy, while contributing whatever is useful from the files of the HPB project. (A duplicate tape of the English items in the HPB database has been offered to the ASECS-inspired, trans-Atlantic project.)

The HPB Project

The HPB (Hand Printed Books) project, which in my last lecture I promised to describe in some detail, has developed within the historical context of short-title cataloging. A history of the project will be found in Chapter I of the HPB Project: Phase III, and as the enterprise is still evolving as an open-ended system for use in teaching and research, I shall attempt to describe here only those elements which seem to be well-established. By using case-study methods in subsequent lectures, I hope to suggest possible ways in which the project could develop in the future, and will confine my remarks in this lecture to the status quo, and to international implications. For the HPB project, though biased by the English-language context and the geographical context (New Zealand, Australia and Canada) within which it has been developed, is not confined to English materials, and, hopefully, will not be confined geographically.

The machine-readable records of the project consisted in 1976 of:

1. a main file of STC surrogates for upwards of 30,000 bibliographically distinct volumes. The STC entries represent several stages in the development of the project's STC cataloging rules, and cover a wide

variety of library holdings.

2. a battery of "collocation files", of which the HPB/NUC Index is the major control file for the main file; others represent individualized access devices to the main file and the HPB/NUC Index.

3. the "GOTHIC" retrieval system for general access to the main file, and for specialized access (using the various collocation files) to parts of the main file.

Most of the entries in the main file are for English publications, but a substantial proportion of the entries are surrogates for material in other modern languages. To provide a copy of relevant parts of the database to the "18th century English STC" project for whatever use the organizers wish to make of it, will require a research experiment for distinguishing by machine between material within the scope of that project and material outside it. The excluded material could make an interesting nucleus for developing STC projects in other languages and cultural areas of the world.

At the British Library meeting, a strong plea was made on behalf of potential users of the 18th century STC, that the bibliographical coverage should comprise "the output of all British presses". Because the users were assumed to be British and American, or people interested in British and American culture this plea met with general approval.

However, if we try to see bibliographical coverage of a multi-lingual, multi-cultural 18th century STC in terms of practical national responsibilities, it seems obvious to me that this 'cultural' partitioning is in fact very difficult to make without recourse to highly-sophisticated bibliographical experts. It is fairly easy for an unsophisticated, intelligent worker on an STC project to determine the language that a book is written in, and the very small proportion of items where decision is difficult would certainly be more easily determined than the often hidden cases of false imprints or of misleading imprints. Those imprints were meant to fool the unwary; difficulties of polyglot books and indeterminate languages are mostly accidental and self-evident. Therefore, I would suggest that an English STC should be confined to the English language, and that the products of British presses in other languages should be included in other language STCs, with appropriate indexes made by the experts who would indeed be better employed in using their talents in creating such an index.

We may thus return to Ranganathan's concept of the characteristics which might form the basis for restricting an "omnibus universal documentation list" to useable lists, and attempt to modify the first of them - Area of origin of the documents. Linguistic nationalism seems to me to be a better base than cultural nationalism to define an area of origin for a documentation list requiring international cooperation. I would even go so far as to say that separate files for books in Latin, Greek, and Hebrew (the common cultural heritage of Western European nations) should be made for 18th century books so as to reduce the size of the undertaking for any particular group of nations whose prime concern is a modern European language. In this

respect, I propose to develop, over the next year or two, the nucleii for a French language STC, a Spanish language STC, and a Portuguese language STC from existing records of the HPB project. This implies an extension of geographical coverage to the whole of North, Central, and South America. Even the technical problem of identifying the language of the records in the database without 'tagging' the entries beforehand is a challenging research project. And I want to reemphasize that the HPB project is not a fully operational system; it is primarily a teaching and research tool.

But to return to our description of the HPB database as it now exists. The three sets of machine-readable records mentioned above have been created with several basic assumptions in mind. I hesitate to call them "hypotheses" as the whole system is not a rigorously-controlled empirical investigation, but rather a loosely-organized project for generating smaller, dependent, empirical research projects. Perhaps the most fundamental assumption is that any system of bibliographical control of early books should be designed to meet the needs of scholars in the humanities, whose characteristic use of bibliographical data differs considerably from that of social scientists and natural scientists. I know of no empirical investigation that has even attempted to establish what the differences are, but I am convinced from life-long experience that the needs of scholars in the humanities are more difficult to categorize or to generalize about than their scientific counterparts. Their approach to 'subjects' (and hence their use of subject classification) is especially remarkable.

A great deal of scholarly activity in a general subject area is dedicated to redefining accepted subject divisions and providing new points of view on old data by rediscovering data not normally associated with the accepted subject area. Discovering new connections between known facts, discovering new facts to call in question established interpretations, and reinterpreting old facts is an activity that on the whole tends to undermine accepted subject classifications, so that general subject classification schemes tend to be of less help to original scholarship than is generally believed by librarians. Special Collection Librarians are well aware that most of the collections in their care were built on principles that are not usually reflected in general classification schemes, or are out of proportion within such schemes.[1] An "author" collection requires a more detailed classification scheme than is normally to be found in a general classification scheme; a "form" collection (such as "courtesy books" or "dictionaries and encyclopedias") is not likely to need subject classification of the general kind, and examples of book production of a particular era or locality will not need a priori decisions about their subject interconnections. Chronology is likely to be more important.

The implications of this major assumption for the HPB project is that a multiplicity of access points to bibliographical data (other than by subject

---

1. I would even go so far as to assert that the latest edition of Lee Ash's Subject Collections, despite its commendable increase in comprehensive coverage, is in fact less useful to specialized scholarship simply because of its attempt to subsume special collections within a general classification scheme of admittedly proved use to the general user.

classification) must be provided.

Another assumption is that scholars are constantly returning to material that is already known to exist, rather than to newly-produced literature, and that much of it has already been described in some detail in the bibliographical literature of the past, and need not be more than minimally described in a bibliographical control system. What is needed is to bring existing descriptions into convenient relationship with the minimal description when needed.

Thus, comprehensiveness of coverage in the descriptive element of the system, and flexibility for manipulation of existing data in the control element of the system become the major desiderata.

Another assumption is that most of the items described in the main file of the project will be housed in Special Collections with peculiar strengths and weaknesses, and that guides to the real strengths of such collections will be an important intermediary function of a bibliographical control system. Thus, although the primary purpose of the control system is to optimize the individualistic requirements of particular scholars in the humanities, it should also be assumed that the control devices should be able to optimize the work of Special Collections Librarians seeking to provide special kinds of intellectual access to their collections. Only in third place will come more general means of access to rare materials.

Bibliographical Control Devices of the HPB Project

Because of these assumptions (and many more like them), a trade-off between simplicity, economy, and speed of compilation in the main file and complexity, individuality, and elaborateness in the control files is being constantly attempted. The main file should, eventually, become a stable, essentially simple, and hopefully accurate set of descriptive records requiring no further modification except the addition of new records.

The control files will always require modification of data as they respond to new information, new relationships, and new uses. They will probably fall into three categories, each with increasing flexibility and decreasing need for permanency. The first category will be a generalized control system, highly structured, designed for a multitude of common requirements for access to the main file. The HPB/NUC Index is being developed at the moment as a possible prototype. The second category will be control systems for particular special collections that will not serve a useful purpose in another collection. Indexes to specialized material (such as dictionaries, library marks, or books containing subscription lists[1]) in our own Special Collections library at the School of Library and Information

---

1. This will ultimately become a contribution to the very interesting project being undertaken at the University of Newcastle by P.J. Wallis and F.J.G. Robinson. The project is described in many places, but The Library 5th Series, XXIX, 1974, pp. 273-5 might be the most convenient and informative. The project is a very fine example of macroscopic bibliographical control.

Science, University of Western Ontario, are being developed at the moment. The third category - the major one - are control systems to aid individual scholars. Some possibilities will appear among the exemplifications of scholarly behavior in the next 3 lectures. The three categories of control system all make use of the basic element which we call "collocation" files.

Collocation Files

As explained in Chapter 3 of the HPB Project: Phase III, collocation files so far developed can be divided into 5 "kinds".

1. the 'perfectible bibliography'
2. the 'substitute for discourse'
3. the matching
4. the 'register'
5. the 'cross-reference'

A particular collocation file (such as the HPB/NUC Index, which is the main control for the HPB project) may be a combination of all or any of these different kinds of file.

It is not my purpose here today to describe in any detail the ways in which collocation files can be constructed and used, but it is perhaps important to emphasize the extreme flexibility and individuality of the complex system of files that can be evolved for use with the main database. Each individual scholar may design his or her own files for his or her own particular purposes; each individual special collections librarian may devise files to provide particular kinds of access to his or her own particular collection; and files such as the HPB/NUC Index may be devised to provide more general kinds of access to a large number of special collections. I shall illustrate some of the possibilities in my next lecture by a case study, and confine my explanation of the 5 'kinds' to a rudimentary explication.

Many bibliographies in conventional printed form provide addresses (usually in the form of numbers or letters plus numbers) to the entries contained in them. A "perfectible bibliography" kind of collocation file then, is a simple list (in any order desired) of selected addresses in machine-readable form which can be added to or subtracted from according to the purpose served by the list. If the items in the original bibliography have been converted into machine-readable form (together with the addresses), the list can be 'passed' over the database and the appropriate bibliographical entries printed out in the predetermined order.

If the database contains all the items, but with a different address system, a 'matching' kind of file consisting of all the addresses in the manually-produced bibliography with (after an "=" sign) the equivalent address or addresses in the machine-readable catalog can be constructed to enable the

'perfectible bibliography' kind of collocation file to be used to print out the desired bibliographical listing. You will note that changes can be made in the 'perfectible bibliography' listing without changing either the matching file or the main database.

A printout from a 'perfectible bibliography' collocation file will be simply a list of bibliographical entries. But scholarly comments on each or some of the entries, explanations of relations between the entries, or any other kind of editorial "connecting tissue" may be inserted into the simplest kind of collocation file and thus, when passed over the database (with or without a 'matching file' mechanism) will print out a "substitute for discourse". Thus, our second kind of file is an elaboration of our first, and different in degree of explicitness. Our third kind is truly different in kind. Essentially, the 'substitute for discourse' kind of file permits individual users of the database to provide themselves with bibliographical data in a desired order with desired annotation for their particular purposes. And it is 'perfectible',i.e., it can be sophisticated as new interpretations of the data, or new data come to hand.

The fourth and fifth kinds, as I shall explain shortly, are only two of many other possible uses to which the system of collocation files can be put. In the first three kinds, we use a slanting bar (/) to indicate the delimiters of data elements within a collocation file, and an "equals" sign (=) is used both as a substitute for a delimiter, and an indication (to the machine) that what precedes and what follows (back to and up to the adjacent slanting bars) are the two parts of an entry in a matching file. We use a plus sign (+) also as a substitute for a delimiter and as an indication to the machine that what precedes and what follows must be printed out as two bibliographical entries with the word "AND" between them. By these means, for instance, a book subsumed in one entry in a particular bibliography, but separated into two or more entries in another, may be accurately matched. Where there are special ambiguities, a machine-readable 'natural language' comment may be added to the collocation file between slanting bars.

These conventions permit further elaborations, and the fourth and fifth kinds of collocation file are the beginnings of such elaboration. A "less than" sign (<) is used to indicate that what follows is part of (or subsumed in) what precedes and a "greater than" sign (>) is used to indicate that what precedes is part of (or subsumed in) what follows. Again, the slanting bar is the delimiter for both what follows and what precedes. The particular application of these two kinds of file in creating our HPB/NUC Index (i.e., the control file for our HPB project) is as a 'register of locations' of particular copies of an item whose address precedes the "less than" sign, and as a cross-reference from one address to another for analytics, misfiled entries and duplicates, etc. So I hope if I give you some examples from the HPB/NUC Index, the principles upon which it is built and the use of such a file will be clear. Let me illustrate:

```
M0606255 782=1089XX<CaJTU
M0606263 791=0385BT<CaBVAU CaDHM
M0606268 795=MI0117<CaDHM CaDTU CaDDCC CaDLUI
M0606270 799=2533UA<CaAEU
M0606270* 05=0426BT<CaDLUI
M0606410 743=MI0123<CaDHM
M0606654 725=MI0263<CaDHM
M0606668* 05=2029MC<CaDHM
M0606669 752=      <CaDKQ
M0606682 760=MI0279<CaAEU CaDHM CaNBSAM
M0606780 730=1208XX<CaNSHK
M0606829 651=1297XX<CaAEU
M0606878 654=1295XX<CaDHM
M0606912 796=MI0336<CaDHM CaDKQ CaDLUI CaDTU CaBVAU
```

```
M0605728 743=MI0180+1179XX/
M0605729 743>M0605728 743/
M0605730 748=1186XX/
M0605730* 05=      /
M0605730* 07=      /
M0605730* 09=      /
M0605731  00>M0605730* 05/
M0605732 749=1246XX/
M0605734 753=1253XX/
M0605734* 05=0120BM/WHICH INCLUDES/0069BM/
M0605735 755=0069BM/
M0605735* 05=      /
M0605738 757=1185XX+1269XX/
M0605739 758=1251XX+1252XX/
M0605739* 03=1086XX/
```

In the first example, we have a combination of a matching file (NUC address before the = sign, HPB address after) and a locations register (Canadian locations of the books) after the < sign. In the second example, we have an example of the matching file and a cross-referencing file, with some of the possible elaborations. Those entries with an * in the ninth column of the NUC address are additions to NUC. The numbers in the 11th and 12th columns are filing devices. "00" indicates that an actual cross-reference occurs in NUC; the other numbers are arbitrary guesses designed to allow room for all possible number of entries missing from NUC to be added later. The blank spaces before the slanting bars indicate that HPB entries have not yet been made and entered in the main file.

I realize that this level of detail is difficult to follow in an oral presentation, but the unavoidable detail should be sufficient to demonstrate

1. that the basic concept of a "collocation file" may be elaborated for a large variety of control devices or combination of them;

2. that a scholar may make a choice of one of these devices or combination of them, or even elaborate one of his own;

3. that all collocation files are independent of each other and, therefore, need not be altered to accommodate alterations in other files;

4. that the address system of the main file upon which the control file operates must in fact not be altered, even though the bibliographical data may be modified.

The Nature of the STC Entry

It should be well understood that the collocation files are used to control the bibliographical data in the main file, and link it with bibliographical data in printed or published form, and that the contents of each collocation file are the result of an act of appraisal. The STC entry, on the other hand, is a neutrally descriptive item which results from avoiding as much as possible any act of interpretation or subjective judgment on the part of the cataloger. The aim is to identify an item of the bibliographical universe by selecting, as objectively as is humanly possible, the minimum amount of data from the title page and the "bibliographically distinct volume". For those of you who are interested in the matter, the data is divided into 14 fields as the following example of an input form indicates.

Input for HPB Main File

AD 1251XX
PA MILTON, [John]
SA trans [Nicolas Francois DUPRE DE ST MAUR]
TA
QA
ST Le paradis perdu
ED New edn
VO [3 vols] Vols I-II
PP Paris,
IM Chez Gianeau
DA 1758
FO 12mo 345, 358pp
NO
PO

The field tags are mnemonic (in English). AD (=Address) is for the nonsignificant but unique HPB 6-character address (1251XX is in fact the 1251st item in a series of entries made from xerox copy of title-pages supplied by contributing librarians). PA, SA, TA, and QA are primary, secondary etc. authors, ST the short title, ED and VO edition and volume information. PP, IM, and DA are place of publication, imprint, and date. FO is format (including minimal pagination information) and NO is for notes for resolving any ambiguities in the previous fields, or for additional bibliographical information needed for identification of the item. PO ("Points") is for points of difference for separating editions otherwise subsumed under the data given before. The precise cataloging rules are to be found in Appendix K of The HPB Project: Phase III, and the principles and rationalization for them are to be found in the The HPB Project: Phase II, Chapter 1. But most of the conventions should be self-evident from the example given above. Letter-by-letter transcription is used wherever possible, but some fields contain data in standardized form.

As can be seen from the example given, the language of the catalog entry is English wherever the data element consists of conventionalized (rather than transcribed) data. This, of course, is true of most catalogs and bibliographies, but it does pose a problem for an international approach to short-title cataloging. If notes (NO) or points of difference (PO) were given in the language that the volume is written in (which seems a reasonable procedure) cataloging could be done only by catalogers fully expert in that language. A more reasonable approach might be to retain the English language in the database and, if necessary, devise some method of machine-translation for the very limited vocabulary used so that the display form of the data for a particular language group of users could be printed out for their convenience. This might be an intriguing research project for the future international development of the HPB project.

Spanish, French, and Portuguese STCs

In order to find solutions to such problems and to explore the implications of multi-lingual development of the HPB project, the three major languages of the Americas will be given priority in the near future. As Spanish is the second-largest language group in the hemisphere, this will be the language of highest priority, even though the Canadian base of the project will ensure that French is not neglected. But we might take Spanish as an example for present purposes to speculate on the ways in which the existing system can be used to extend bibliographical control beyond the English language.

First of all, the main file of STC entries must be systematically augmented by cataloging available books in the Spanish language. Particular authors (e.g., Cervantes) will be treated in similar fashion to John Milton, who was used as a model for the development of the system, as reported in The HPB Project: Phase III; particular Special Collections containing Spanish-language materials befor 1801 will also be cataloged in similar fashion to the New Zealand, Australian, and Canadian collections incorporated into the database as reported in The HPB Project: Phase II or in the modified fashion (using the National Union Catalog: Pre-1956 Imprints as a selection device) reported in The HPB Project: Phase III. The latter set of procedures will

of course also be part of the development of the HPB/NUC Index as a control system for the main file.

Most useful, however, will be the development of general, special, and individual collocation files for the control of bibliographical data both in the main file and in selected bibliographical tools in conventional printed form. Selection of the tools to be incorporated into matching files and 'perfectible bibliography' files will depend to a great degree on whether each item in the bibliographical instrument has a unique address either given or capable of being inferred. High priority will thus go to incorporating the work of the great Chilean bibliographer José Toribio Medina. His Biblioteca Hispano-Americana contains over 6000 pre-1801 items in chronological order, and a matching file linking his entries with equivalences in NUC: Pre-1956 Imprints would form an excellent nucleus for a 'perfectible bibliography'. Subsequent additions, corrections and deletions to Medina could be incorporated quite easily. But we must remember that Medina himself once said:

> He trabajado mucho y me he cansado poco.

It will take a Herculean hero like Medina himself to digest into machine-readable form the enormous amount of bibliographical work that he performed for Spanish America during his lifetime, and that hero will need to be prepared to exercise a similarly untiring use of supplementary catalogs and bibliographies to improve upon his work.

Medina, like many bibliographers whose understanding of the purpose of their work was to serve humanist scholars, presented the results of their scholarship in chronological order, which necessitates supplementary indexes. As the order in which machine-readable entries are kept on tape or disk is irrelevant to the order in which the entries may be displayed, we may use Medina's order as only one of many possible collocation files for ordering the data. But at this point, it might be as well to re-emphasize the point that subject classification, and author-title order, is often thought by scholars to be of less use than chronological order. It seems fitting that another bibliographer, Lawrence C. Wroth, in his appraisal of Medina's work felt obliged to quote his own opinion from Standards of Bibliographical Description, 1949, on 'a sort of cosmic orderliness about chronology' and also felt obliged to reprint his forthright statement:

> No instrument more effective for the fragmentation of thought could be devised...than the [bibliographical] catalogue based upon the alphabetical principle, and, after all, one of the services of bibliography is the organizing of thought for the purposes of synthesis.

Attempting to match Medina against NUC will certainly demonstrate the difficulties, but his "Indice Alfabetico" will undoubtedly simplify the problems of competing systems of arrangement. Other kinds of indexing in the form of collocation files will then become possible once Medina's works have been converted into machine-readable form. But let me leave the Spanish American development of the HPB project, and make some generalizations about collocation files.

Collocation Files and Existing Bibliographies

If subject classification of the items (or a group of items) in the main file was thought to be desirable, collocation files for each subject could be compiled. In a particular Special Collection devoted to a special subject, an existing subject bibliography could be identified which was thought to be the best classified and most comprehensive in coverage (it could be a library catalog!) and the items in the bibliography matched (by machine-retrieval) with the items in the collection. Those items in the bibliography NOT found in the collection could be relegated to a separate collocation file to be used as an acquisition tool. Those items in the collection not found in the bibliography could be added at the appropriate point in the classification scheme and the address added to the collocation file. Any new item acquired could also be classified and added to the collocation file. In this way, a "perfectible" catalog, classified according to an existing scheme would result.

But if a special classification scheme was desired, the collocation file could also be used to embody the decisions made as to where particular items should appear. They may in fact appear in more than one place (with annotations where necessary), thus overcoming the limitations of library classification schemes that govern the shelving of the book as well as its subject content. Thus, although one of my persistent themes in subsequent lectures will be a plea for the use of arrangement schemes and classification schemes (other than subject classification) as devices for bibliographical control, the HPB project does not preclude the possibility of the kind of universal subject classification approach so strongly advocated by Dr. Ranganathan.

Lecture 3

BIBLIOGRAPHICAL CONTROL OF A PRODUCER OF DOCUMENTATION

I hope that the principles outlined in the previous two lectures and the brief desciption of my experimental project for developing a means for macroscopic bibliographical control of early books will have provided you with a sufficiently understandable theoretical base to enable me to illustrate with some practical examples the application of those principles. These examples will help me to reinforce, and perhaps restate, those principles and establish their validity. Because I am giving these lectures in India, I thought it might be appropriate to choose, as my field for exemplification, the units of the bibliographical universe produced by the man reputed to be the first letterpress printer in India - Henry Hills.

Henry Hills in India

J.B. Primrose, over 30 years ago[1], first pointed out that Henry Hills was employed (at a salary of ₹ 50 per annum) presumably from May 15, 1674 to May 15, 1678 as a printer to the East India Company in Bombay. Because there was so little use for his professional expertise, he was (somewhat unsatisfactorily) employed part of the time as Steward of the Factory at Surat. We have evidence that he took a printing press, paper, and types (English letters!) to Bombay or Surat, and that Bhinji Parekh, the Company's broker, arranged to have types made, probably in the Devanagari character, by Indian metalworkers, in an attempt to use Hills' services properly. We have no "bibliographical control" of the papers that were printed in "the Banian Character" in 1676 nor of any printed in the English character, simply because nothing of Hills' printing from the 1674-1678 period seems to have survived. Bhinji attempted to obtain from England a type-caster in 1677, but Hills was sent back to England in January 1678/9 under something of a cloud, seemingly caused by his financial problems and by dissatisfaction with his work as printer and steward. No further attempt at printing in Bombay took place for over a century.

Despite this rather regrettable beginning to his career, I believe that Hills' total career, which has never been given adequate treatment, would make a good subject for research, and would illustrate well the complex needs of scholars for different kinds of bibliographical control of contemporaneous documents. But from the point of view of my Indian listeners, the reason for undertaking research into the later life of Henry Hills might be more cogently stated to a desire to review Primrose's interpretations of the evidence he found for Hills' visit to India. To put the matter in more generalized research terms, the scholar will treat Primrose's conclusions sceptically and ask him (or herself) certain questions that

---

1. J.B. Primrose, "A London Printer's Visit to India in the Seventeenth Century", The Library, 4th Series, XX (1939-40), pp. 100-4.

might be developed into hypotheses for further testing. Was the failure to establish printing in India in the 17th century caused primarily by Hills' incompetence in financial, printing, and work habits, as implied by Primrose? What kind of man was he? In the absence of new sources of information for the 1674-1679 period, can his post-1679 career throw light on his probable behavior in India? I shall not, of course, attempt to answer these questions, or even develop appropriate hypotheses in this series of lectures. My purpose is to investigate how a scholar can be provided with (or provide for himself) access to bibliographical data that will help him to solve such problems.

Primrose's account of Hills in India was based primarily upon type-written copies ("exemplars") of correspondence ("texts") carried on between the Governors of the East India Company in London and their representatives in Surat and Bombay. I suspect that an index of names rather than a subject index would provide the scholar with optimal content access to such documentation in the India Office. A second group of papers of use to Primrose were printed Court Minutes, a 1935 "exemplar" of the original 1674-76 "texts". Access to the contents of this publication is in fact mainly by nominal index, which contains a few index entries for institutions, and for substances such as goods that formed the staple of East Indies trade.

Primrose's rather sketchy and inaccurate biographical information about Henry Hills and his father of the same name comes from what scholars call "secondary" sources - biographical dictionaries (accessed by name)[1] and from either the few references or the sole periodical article[2] quoted in one of those dictionaries (Primrose thus used a crude form of citation index). Primrose's sources tend to emphasize Hill's activities in the second decade of the 18th century as a 'pirate' printer of poems and sermons. How would a modern scholar go about obtaining better and more substantial biographical information about Hills, and what bibliographical tools would he need to exercise a maximum of bibliographical control over his materials?

Bibliographical Desiderata for Research into Henry Hills

First and foremost he would want a list of all publications in the production of which Henry Hills was involved. Second, he would want a list of all primary materials that mention Hills or his activities. And, perhaps third, he would like a list of all secondary accounts of Henry Hills. But the irony of scholarly research is that such lists are produced (usually) only by the scholar who needs them. He or she puts his or her lists together by using a multitude of bibliographical tools produced for other (usually more general) purposes.

Such tools as do exist are characteristically found through subject bibliographies, however. These subject bibliographies are usually limited

---

1. For those interested they are Plomer's Dictionary of Printers and Booksellers, 1668-1725 and The Dictionary of National Biography.

2. Edward Solly's "Henry Hills, the Pirate Printer", The Antiquary, XI, 1885, 151-154.

according to all 6 of the characteristics we talked about in the first lecture, and vary so much in comprehensiveness or selectivity, generality or specificity, relevance, size, and availability that bibliographers have in the past been forced to produce "bibliographies of bibliographies" (usually arranged according to subject classifications) to enable scholars to choose the right selection of interdependent subject bibliographies. Part of the machinery of control of these bibliographies of bibliographies has been annotation (according to presumed purposes of the users). Such annotation has occasionally been elaborated into books about bibliographies of bibliographies that contain (either in footnotes or an appendix) a bibliography (select or comprehensive, annotated or classified) of bibliographies of bibliographies. The most comprehensive example of these is Archer Taylor's A History of Bibliographies of Bibliographies, 1955. Contemporaneous bibliographies tend to be overlooked or be judged 'superseded' by many compilers of bibliographies of bibliographies, so the scholar using early books tends also to work outside this structure of bibliographical aids, or to assess the limitations of those aids very carefully. These limitations have a direct bearing on the relationship between the scholar's specific purposes and those general purposes presumed by the compilers of such tools.

Let us do a "dry run" through the hierarchical structure to illustrate. We want a bibliography of bibliographies that will lead us to the most useful of bibliographies for our purpose - the elucidation of Henry Hills' activities in India from 1674-1678. We might guess that we need bibliographies dealing with place (India and England), a particular kind of enterprise (printing and publishing), a particular time (Hills' lifetime, approximately mid-17th century to 1713), and a particular man (Henry Hills). Obviously, the key subject classification is "printing and publishing", and the key form classification is "biography". But we may have to modify our preconceived notions about subject headings when we come to consult the actual bibliographical instruments.

Archer Taylor's history gives us a "bird's-eye" view of the field of bibliographies of bibliographies, and his fairly short text provides an admirable guide to help us select for our present purposes the appropriate bibliography of bibliographies from the bibliography in the back of his book. His index helps us to review our decision, but there can be little doubt that Theodore Besterman's World Bibliography of Bibliographies[1] is the most likely to yield the information we want. Taylor warns us of the weaknesses and limitations of the work, and in particular criticizes the alphabetical arrangement of many small headings which makes the book difficult to use efficiently. Despite the elaboration of cross-references in the 3rd and 4th editions, this is as true of the 1965-6 publication as it was of the 2nd edition of 1947-9 that Taylor analyzed with such care. In the preface to the fourth volume of the 4th edition (1965-6), Besterman reports that his 5-volume book contains 117,187 separately-collated volumes of bibliography classified under 15,829 headings and subheadings. (It may be of interest to note that 55,373 of the volumes are in the English language -

---

1. The fourth edition in 5 volumes was published in 1965-66 by the Societas Bibliographica, Lausanne, Switzerland.

47% of the total, with German [16%] and French [14%] next in importance.) Our first task will be to discover which of those many thousands of headings and subheadings will yield references to bibliographies containing either lists of Hills' publications or lists of primary or secondary sources for the study of Hills and his work.

Any hope that Hills' name would appear as one of the headings or would be mentioned in the Index volume is soon dashed. In his preface, Besterman tells us that "bibliographies of the productions of specific printers and presses are not entered under their names but only under the heading Printing". In an attempt to satisfy our scholar's first need, we might hasten to consult this heading. The introductory note gives cross-references to other headings, of which Bibliography and Forged Books seems hopeful. But the entry itself has subheadings that help us home in on the two subheadings most likely to yield guidance - the countries India and Great Britian. Nothing relevant to Hills comes from a search under the latter, but under "India" we find 3 very rare bibliographies of early Indian imprints, two by D.C. McMurtrie, and one by K.S. Diehl, so rare that I could not conveniently consult a copy of each when preparing this lecture. I doubt, however, whether Hills is mentioned in them.

In attempting to satisfy our scholar's second and third needs (a listing of primary and secondary sources about Hills and his career), we need to understand other principles upon which Besterman's compilation is made. His preface, apart from a critical review of his 3 major predecessors (Petzholdt, Vallée, and Stein) and dismissal of other minor ones, consists of an account of what is included and excluded from consideration, and a strong defence of his reasons for choosing a subject classification and an author index in preference to other arrangement schemes. On scope, the most important points are that his compilation is strictly limited to separately-published bibliographies (defined as "having separate pagination") and "true" bibliographies (defined as "a list of books arranged according to some permanent principle"). The first limitation is a serious one for our purposes, as our man is not of such social or historical importance that a separately-paginated "true" bibliography is likely to exist. A list forming part of a scholarly article, one included in or appended to a dissertation on a broader topic, or even a list arranged according to an ephemeral principle is more likely to exist. Besterman asks for "common sense" in interpreting "permanent principle", but he uses it to exclude bookseller's catalogs, general library catalogs, and several other lists from him compilation. We will have to seek all these kinds of instruments of bibliographical control (and many others mentioned in the preface) outside the hierarchical scheme.

But within these limitations of scope, we must come to grips with Besterman's classification scheme. The best principle to follow is to seek the specific heading before the general heading. We might try East India Company before India, for instance, especially as Bestermen is explicit in claiming that he prefers Subject before Country. ("Subject" seems to include "Form" and other headings in Besterman's rather loose use of the term.) On one page of the preface he gives a list of subject headings which are subdivided by country, and a quick scanning suggests that we ought to consult Biography, Broadsides and Pamphlets, Drama and the Stage, Poetry, Printing, and Sermons.

(Note the number of form headings!) He admits, however, that some subjects must be subsumed under a country heading, and lists those for Great Britain in illustration. A search for "East India Company" merely produces "East Indies" as the nearest alphabetical heading. References to Primrose's printed sources are to be found there, but most of the bibliographies are for Dutch material for the Netherlands East India Company.

Consultation of the "India" and "Great Britain" sub-headings for the six broad headings selected above yields a mixed bag of general bibliographies that might lead us to primary or secondary sources about Hills. For instance, "Biography" suggests Phyllis Riches' Analytical Bibliography of Universal Collected Biography, 1934, but Hills' name does not appear among the 56,000 entries. She lists 123 names in her subject index under "Publishers, Printers, and Booksellers". An older bibliography, Edward Edwards and Charles Hole, A hand-book to the literature of Collective Biography, 1885 is not easily consulted, but as it contains only 100 entries, it is unlikely to be of help. (Besterman's alternative to annotation, an estimate of the number of entries in each bibliography, is a highly useful device for bibliographical control.) Finally, the periodical publication of the H.W. Wilson Company, Biography Index (1946-) yields nothing on Hills. Cross-reference from the introduction to a subheading of "Bibliography" and to "Family History" does not yield anything relevant to our needs.

However, it is interesting to recall that Hills himself classified his published output as All Sorts of Sermons, Poems, and Plays, &c in several of his own advertisements, so consultation of three of the Besterman headings should yield material of interest. Under "Sermons" appears a 325-item bibliography by William Hawes which claims (in 1709) to be A compleat collection of all the sermons that are printed, and sold for one penny, two-pence, or three-pence.... This is a bookseller's catalog that fortunately escaped Besterman's exclusion policy. But as it does not specify the printers or publishers of the sermons, it does not help us identify Hills' output. Sampson Letsome's An index to the sermons, published since the Restoration (1734, 2nd edn 1751) and The Preacher's Assistant (1753) are very valuable tools however. The index consists of 6 columns; the first two give biblical texts by Chapter and Verse in the normal canonical order, the third gives the authors of sermons on that text, the fourth designates which volume in a multi-volume set the sermon occurs in, the fifth gives edition (e.g., "fol" or "Clarke, 17 Serm") and occasion (e.g., Tr.S. = Trinity Sunday), and the sixth column gives page number (if in collected sermons) or S.S. = single sermon. No publishers are given, but use of the index after Hills' output has been determined would be essential for study of the reasons for his publishing those sermons and not others. Similarly useful tools are revealed under the other headings, but nothing specifically to our purpose.

In fact, to cut a long story short, the most useful bibliographies that a careful search of Besterman produces are to be found under "India" (the subheading "History" lists 8 guides to, press lists of, or reports on, the records in the Record department of the India Office) and "English Literature". Under the latter subject heading appears Bateson's Cambridge Bibliography of English Literature, 1940, which yields (in Volume II, devoted to the period 1660-1800), an entry for Edward Solly's 1885 article on Henry Hills which was Primrose's major source of biographical information.

Strangely enough, the 1971 re-edited Volume II does not list this article, but lists two contemporaneous pamphlets on Henry Hills Senior, Primrose's 1940 article, one that I wrote in 1960, and by far the most significant secondary source up to time of compilation, R.P. Bond's article in The Library in 1963. Using Bond as our take-off point, we could list all the sources he cites as the nucleus for the second of our scholar's bibliographical desiderata list.

However, under Besterman's "English Literature" heading, we can find periodical bibliographies (under 3, Writings on English Literature) that hopefully might extend the CBEL listing for the third of our scholar's needs. In particular, English Literature 1660-1800 A bibliography of modern studies compiled for Philological Quarterly, a periodical bibliography beginning in 1926, yields a 1958 article which I wrote, and a very fine article by John W. Velz printed in Publications of the Bibliographical Society of America in 1969, as well as Bond's article. This good yield is a result of a highly-specialized section in the bibliography devoted to "Printing and Bibliographical Studies". It is a pity that periodical bibliographies in the field of printing studies have developed too late for inclusion under Besterman's "Bibliography" heading. But he should at least have listed the two cumulations of the annual Studies in Bibliography lists. The first was published as "Selective Check List of Bibliographical Scholarship" in 1957 and fits Besterman's criteria; the second was published in 1966 (too late for inclusion). Neither, unfortunately, lists anything for Hills. The annual list was terminated in 1973 with the decision to join ABHB (Annual Bibliography of the History of the Printed Book and Libraries), an international attempt by the International Federation of Library Associations which is gradually gaining national cooperation and thus comprehensive bibliographical control of the field. Nothing on Hills has yet appeared in the journal, however. The 4th volume (1973, published 1975) contains 3400 entries and added entries. The classification was borrowed from Bibliography in Britian (1962-), itself a parallel publication to the Studies in Bibliography Check List. It fails to register items on Hills also.

Systematic searching (I have not done more than suggest the detailed procedures) thus throws up a selective list of secondary sources of information about Hills, but only the more scholarly items will in fact provide us with a list of primary sources. Velz managed to operate very well without knowledge of either my articles or Bond's, and probably did better work by concentrating on primary sources such as The Calendar of State Papers (Domestic) than upon the primary sources used either by me or by Bond. Subject bibliography is of only limited use in satisfying our scholar's third need, and the citations in the secondary sources still need to be augmented by new sources reached usually by imaginative guess-work about the kinds of document or printed material that ought to exist and may perhaps have survived. Let us, therefore, concentrate on the question: How could we go about utilizing macroscopic tools to satisfy at least the first need of our scholar?

Compiling a Bibliography of Hills' Publications

"Imaginative guess-work" will suggest that Hills himself must have advertised

his publications in various ways, and that such booklists (despite their not being organized on a "permanent principle") would form a valuable source of information for the compiler of a bibliography. My first attempt to cope with getting such material under bibliographical control was reported in an article published in New Zealand in January 1974.[1] A similar kind of study might be necessary for Henry Hills, Senior. A cursory search of some of the publications of Henry Hills Senior during 1688 (the last year of his career as King's Printer) revealed a 2-page booklist in the 4th edition of Richard Hudleston's A Short and Plain Way 1688 (Wing H3257) and a 3-page booklist in the 12mo edition of the same date (Not in Wing). Most of the items are Catholic propaganda pieces, whereas Henry Hills Junior was proud of his Protestant printing (in particular the notorious The Prince of Orange His Third Declaration). This bears out Velz' vivid account of family rivalries caused by religious divisions.[2] But it also demonstrates the importance of such neglected publication lists as potential macroscopic tools for bibliographical control.

A second kind of macroscopic tool that might be put to use is the short-title catalog as described in my previous lecture. Fortunately we may use Wing's STC as a source of information about the publications of both Henry Hills Senior and Henry Hills Junior during the lifetime of the older man, who ceased his printing career on December 12, 1688.[3] The use of this manually-produced STC will suggest ways in which a machine-readable STC might be made even more useful. In addition, Morison's Index of printers, booksellers, etc. in Wing[4] lists the contents of imprints for each person in chronological order, so the capability of machine indexing according to two principles is anticipated. We may supplement such analysis by using advertisement lists of Henry Hills Senior to distinguish the work of the two men. The tracking down of such lists remains to be done, however.

Eight Wing items contain "H.H. jun." in the imprint, and 46 contain either "Henry Hills junior" or "H. Hills junior". The scholar's first task is to distinguish which of the 280 or so items with the words "H. Hills" or "Henry Hills" are by the senior man, and which ones by the junior. This is an act of appraisal facilitated very much by the neutral descriptive listing of all the items for study - a selective display capability of the computer that must in the present circumstances be arduously simulated manually.

---

1. William J. Cameron, "An introduction to the bibliographical control of a piratical printer: Henry Hills", Words, Wai-te-ata Studies in Literature, Wellington, New Zealand, January 1974, pp. 67-72.

2. John W. Velz, "Pirate Hills" and the Quartos of Julius Caesar, Papers of the Bibliographical Society of America, LXIII, 1969, pp. 177-193.

3. Plomer, et al, A dictionary...1668 to 1725, Bibliographical Society, 1968, pp. 154-155, first noted this date; Velz gives even more detail and cites original sources in evidence.

4. Velz demonstrates admirably how a scholar can use such an index.

Henry Hills Senior was one of the King's Printers, so 100 items of the 280 can be readily attributed to him. As one of the Kings to whom he was printer was the Catholic James II, much of the semi-official Catholic propaganda by Father Philip Ellis, John Gother, John Johnson, and Louis Sabran with Hills' name in the imprint may also be attributed to him, and he was associated as publisher with Matthew Turner, the Catholic publisher in at least 10 items. Thus, the list of over 280 that Henry Hills Junior may have been concerned with may be reduced to 83.

This list of 83 items provides the scholar with locations, so examination of an original may help him to shorten the list, either on typographical evidence (by comparison with those in the list of 54 certain items) or because some indication of an address may be found in the original imprint. The work of refining the two lists would most helpfully result in a collocation file for use with the HPB files.

However, even without consulting the original copies, the scholar can further appraise Hills' career after his return from India by ordering and re-ordering those Wing entries that can probably be attributed to Hills Junior. The conventional author/title/date filing system of Wing enables the scholar to conjecture that certain items with no printer's name in the imprint may in fact be by Hills - conjectures that will be validated later by comparing the originals of the known and conjectural printings. For instance, the earliest known item printed by Hills is William Sherlock's A resolution of some cases of conscience, 1682/3 (Wing S3336) and a 1683 edition (S3337) also contains Hills' name in the imprint. Both editions were printed for Fincham Gardiner. Later editions (after Gardiner ceased publishing in 1684) were printed for Basset and Tooke and contain no printer's name, but it is possible that Hills was the printer, for he printed at least 3 items for those publishers in 1683, as well as 3 for Gardiner and, in 1684, 2 for Tooke and Gardiner. Conjectural work of this kind may eventually lead the scholar to detect books printed by Hills between his return from India in 1679 and this first known item. On more solid ground, the scholar will also be able to determine that the "H. Hills" who printed the 8th edition of The Laws of Jamaica, 1683, (Wing J124) was indeed the "H.H. jun." who printed the folio edition, 1684 (Wing J125). Both were printed for Charles Harper. There are other examples of ambiguity that can be resolved by comparison of originals.

Classification according to format of known and conjectural items for which Hills was printer or publisher provides the scholar with tentative conclusions about his printing and publishing activities. (The use of format for retrieval purposes was in fact rejected by the designer of the GOTHIC system as unimportant. Prejudging the issue without consultation with scholarly users has its dangers!) Among the folios are Hills' most ambitious work (Aesop's Fables 1687, printed for the illustrator, Francis Barlow; and The Laws of Jamaica, 1684, with a fine fold-out map). Also among the folios is his most ephemeral work, which can be classed with several broadside poems and prose items. By far the greater part of Hills' output before 1689 was the standard quarto tract of four or five sheets, devoted to sermons, moral or religious tracts, or controversial pamphleteering. He even produced a collected edition of moral and religious tracts in 1685 that ran to 3 editions by 1698. The material is standard

Anglican publication for a wide range of respectable Anglican booksellers, and is supplemented by what look like standard reprints (in the usual quarto format for the traditional publishers) of plays by Shakespeare, Crowne, and Lee.

As Velz demonstrates, microscopic bibliographical control of these reprints will enable us to reach much more sophisticated conclusions. But that topic must wait till my final lecture. This solid core of printing on the whole indicates that Hills was just one of many solid citizens in the printing business. A few quarto political tracts and one or two popular accounts of murder and captivity do not call this conclusion into doubt.

The smaller numer of 8vos and 12mos covers a broad spectrum of items typical of jobbing work, but the books mainly consist of political, moral, or religious tracts. Detailed subject analysis of these or of the rest of Hills' output would merely demonstrate that his market was the major quotidian one of the period. Detailed textual analysis may reveal other, more interesting, conclusions.

Consultation of Wing's listings for 1689-1700 seems to suggest a decline in Hills' output, especially as one third of the entries are in fact undated items that were published between 1706 and 1710! The broadsides, pamphlets, and books published during the last decade of the 17th century show also a diminution in the proportion of quarto tracts. But it is possible that Hills' name did not appear in the imprints of the books he published. The need to distinguish his work from Henry Hills Senior had ceased. Velz, however, by using Calendars of State Papers and of Treasury Books, is able to demonstrate that Hills was employed as messenger to the press in the reign of William and Mary "appointed to inspect printing presses for the discovery of unlicensed books, pamphlets, and newspapers" and David Foxon has published a MS which vividly confirms his activities. (See the illustration in Libertine Literature in England 1660-1745 [1965] opposite p. 16.) It is an account rendered by Hills to the Stationer's Company for buying copies of suspect books in March 1687/8.

David Foxon is a bibliographer of much more central interest to our present concerns, however, for he is a pioneer in creating a "useable substitute" in the absence of an 18th century English short-title catalog. He describes his 2-volume publication English Verse 1701-1750 (1975) in his preface as "a short-title catalog with rather more bibliographical sophistication, though necessarily limited in scope". In the light of our discussion of the six categories for limiting useable substitutes for universal bibliography, it seems significant to me that the scope is limited (apart from area of origin and period of production) by form; in this case literary form (verse) with some attention to bibliographical form. His "STC with frills" contains 10,000 entries, described accurately in his sub-title as "a catalogue of separately printed poems", and hundreds of unnumbered entries which are also described accurately in his sub-title as "notes on contemporary collected editions". The bibliographical sophistication (which he describes at some length in his preface and introduction) is to a large extent peculiar to the material which he catalogs, and is not generalizable to other parts of the bibliographical universe even of the same period. But much of it is appropriate to what we have termed "microscopic bibliographical

control". His use of the position of signatures for distinguishing resettings of type is typical of some of the devices used in the "points" field of an HPB entry, but many of his bibliographical notes we would think were more appropriate to one of our peripheral collocation files.

Volume I contains the catalog; Volume II contains a number of indexes which represent to a certain extent alternative classifications. The first one is peculiar to the nature of the material cataloged - it is a first-line index to the poems - and its importance occurred only to the compiler in the course of his researches. (Indeed, the disarming candour with which the compiler describes the evolution of some features of his work is a very helpful characteristic of his preface and introduction. It helps us to understand how much the organization of the work owes to tradition, especially British Library tradition, how much to modern bibliographical practice, and how much to a response to expressed scholarly needs.) The first index is, in effect, the equivalent of an index to individual articles in a periodical. The second index is chronological (subdivided geographically), and is a true alternative to the author/title/chronological arrangement of the main catalog. The third index is to imprints (geographically divided, and subdivided by names and imprint dates). Like the chronological index, this is of very great use to scholars, and is so generalizable that it must become one of the features of a machine system for bibliographical control. The fourth index is an idiosyncratic subject index called "Bibliographical Memorabilia" the purpose of which is probably so tied to the state of bibliographical studies as it exists at the end of the third quarter of the 20th century that its value to bibliographical scholarship will decrease with time. Like the index which follows it, "Index to Descriptive Epithets", its special value is its contribution to the evaluation of tools for bibliographical control.

The last index, called a Subject Index, is in fact a misnomer. It is a Nominal Index and a KWIC (keyword in title and subtitle) Index, with "certain broad subject headings" interspersed. The nominal index distinguishes between subject terms and titled authors by using the title if the nobleman is the subject, and by using his family name if he is an author. The KWIC index includes verse forms (with implicit subject content as dictated by contemporaneous literary theory), and even other form words if they appear in the titles (this is not possible in a true STC, as much of the data is not recorded). A study of the "broad subject headings" and their use by those who consult the book would be of great interest to those of us who want to place subject classification into its right context as a device for bibliographical control.

Each one of these indexes was compiled after the actual catalog was more or less completed, so that the exact lineaments of the indexes were very much influenced by the kind of data which Foxon had gathered. A great deal of the annotation was nevertheless gathered specifically to help make the first three indexes, especially the second and third, an essential part of the access system. If we were to attempt to incorporate such devices in a machine system, a careful analysis of their strengths and weaknesses would help very much in the design stage. For instance, the imprint index is not just an index to names in the imprints (a KWIC system easily reproducible by machine) but is also an index to the additional notes taken from booksellers' advertisements, etc. And the chronological index is not an automatic index to imprint dates, but an index to the preponderance of evidence presented in the annotations. These refinements on automatic

indexing must be built into our HPB system through scholarly use of collocation files.

Sophistication of HPB Collocation Files for Providing a Hills Bibliography

To provide at least some empirical evidence of what refinements of retrieval would be needed to compile from the HPB database the nucleus of the first of our scholar's desiderata (comprehensive listing of an exemplar of all publications in the production of which Hills was involved), I decided to create from even our embryonic database an "editing file" for Henry Hills by retrieval on the name "Hills" in the imprint field, and report frankly to you on the results - "warts and all".

The database at time of retrieval consisted of

1. over 19,000 miscellaneous items from the "first main file"

2. nearly 13,000 miscellaneous items from the "second main file"

3. 451 items from a file compiled as a side-product of actual research work on aspects of Hills' activity.

All three files in the database required addition of data to the PP (place of publication) and FO (format) fields, and select portions are, therefore, being transferred to editing files before being transferred to the new HPB database. Although there was no duplication of entries in (1) and (3), there would almost certainly be duplication in (2). But the selection from each file would enable me to reach some conclusions about what refinements would be necessary in retrieval before the editing file truly represented a nucleus for the desired list.

The 451 items in the third file included 122 items comprising possible sources for Hills' reprints, publications of his piratical rivals, and other contemporary publications of relevance to research into Hills' activities. Comparison with what actually appeared in the editing file would help categorize what items without Hills in the imprint ought to be added. The 13,000 items in the second file include many duplicates with slight variations in cataloging data which might help to determine how easy or hard it would be to eliminate duplicate entries in a macroscopic file. The 19,000 items would probably include many items that were not strictly connected with our Henry Hills and would help determine how to eliminate "noise" in retrieval.

The results of retrieval were largely as expected, but some surprises occurred that helped to prompt analysis and explication.

Of the 451 items, the expected 122 did not appear, nor (unexpectedly) did 74 others. Careful analysis of these 74 items in the original proofing printout of the data file (dating from the time the entries had been added to the database), revealed the following categories of explanation for the omissions:

1. Hills' name in colophon, not imprint: 5 items
2. No imprint: 27 items
3. "H.H." in imprint, not "Hills": 10 items
4. Data deleted at an earlier time: 11 items
5. Inexplicable omissions: 19 items
6. Inexplicable repetitions: 2 items

Obviously, a search of the "NO" field would have avoided the first category of problems. A search of both "IM" and "NO" fields on "H" only would cause a great deal of "noise" as it would supply all names with an "H" for the forename. A sophisticated phrase retrieval system could, of course, be developed, but would almost certainly not be cost-effective.

No machine system is possible for solving the second category of problems. If a scholar's sophisticated decision (based upon typographical or other evidence) is not incorporated into the imprint field, then the entry cannot be retrieved. As explained in The HPB Project: Phase III, pp. 34-36, the addition of such speculative data to the database is minimized in favour of a more useful incorporation of the speculations into a collocation file. The solution to the problem is, therefore, the compilation by a scholar, not the database manager, of a specialized collocation file. This, in effect, is what Foxon was incorporating into his catalog, rather than into his index.

The fact that 11 items that were eliminated during the editing stages of the database have caused our fourth problem should demonstrate why it is important not to keep on editing the data in the database.

The "inexplicable" omissions and repetitions are almost certainly caused by "bugs" in the retrieval system, as most of them occur among the first half-dozen entries and the last 25, and there is a peculiarly regular reversal pattern evident throughout the printout between pairs of entry. The GOTHIC system still needs tinkering with before we can fully trust it.

From the "second main file" part of the database, only 53 items were retrieved. Thirty-two of the 53 were duplicates of items among the ones already discussed, 3 were duplicates of items retrieved from the "first main file" and 1 was a duplicate of another item elsewhere in the second main file. No real difficulty was experienced in detecting these duplicates. The remaining 16 items, we may conclude, ought to remain in the editing file. But at least 8 of them constitute "noise", for they are almost certainly productions of Hills' father, or his "assigns". As Velz points out, Gilham Hills, Henry's younger brother, was his father's successor to his rights as King's Printer, as Henry Senior by will disinherited his anti-Catholic son. Many of the imprints during the lifetime of Henry Hills Junior bore the wording "By the assigns of Thomas Newcomb, and Henry Hills, deceas'd". The "descriptive" listing on the printout thus gives "appraisal" power to the reader. A collocation file incorporating the addresses of the entries rejected as a result of such appraisal is easily made to serve as index to the "editing file", and thus permits the entries to remain for upgrading and editing of the cataloging data prior to transfer to the new database. The collocation

file could be used later as a "stop list" to eliminate Henry Hills Senior from the bibliography we wish to develop for our scholar. Or it could be used to create a bibliography of Henry Hills Senior. The 8 items from the second main file include some with such imperfect cataloging data that possible duplication of definitive identifications are well nigh impossible before the originals have been checked.

However, 167 entries were retrieved from the first main file, and these, on the whole, only require the usual additions to the PP and FO fields before being transferred to the new database. Cursory inspection of the printout of this part of the "editing file" nevertheless suggested that some items in the old database had not been retrieved because of quirks or "bugs" in the retrieval system. Nevertheless, the experiment was conducted in order to analyse the items that were retrieved, but should perhaps not have been retrieved. The greatest number, of course, are the productions of Henry Hills' father (or his assigns). They number only 18, and the addresses are easily transferrable to the collocation file mentioned above. Another Hills, whose forename initial is given as "P" is found in one entry, and a collocation file for him (if he really existed) could also be compiled if necessary. It's probably a misprint in the original imprint or a deliberate effort to mislead.

It will be readily seen, from this account of an actual experiment in retrieval from our still-developing database, and still-developing GOTHIC retrieval system, that the work which Foxon put into his catalog can be incorporated into the HPB system with (probably) less clerical effort and similar but more flexible results. For instance, the collocation file for Henry Hills Senior could be sorted by machine into chronological order, author order, or title order, thus avoiding the toilsome manual labor of consulting card catalogs and emending them. The printout display form, too, enables appraisal by comparison to be carried out more efficiently. But perhaps a more persuasive demonstration can be carried out on a portion of the Henry Hills file that has reached a stage of completeness comparable with that of Foxon's printed bibliography.

Bibliographical Control of Sir William Dawes

No matter how good the macroscopic tools for bibliographical control, nor how many different ones there are that might mend the deficiencies of others, the scholar will find them of limited absolute value when he is concerned with smaller parts of the bibliographical universe which they cover. But this should not be a matter of concern; what should concern the compilers of those tools is that they perform their limited functions well. Let me illustrate from the writer (or rather preacher) whom Henry Hills pirated more often than any other - Sir William Dawes, Bishop of Chester, and later, Archbishop of York.

A collocation file of all separately-published books and pamphlets by Sir William Dawes was created by using the following macroscopic tools and verifying their accuracy against the original exemplars wherever possible:

1. The National Union Catalog: Pre-1956 Imprints

2. Wing's STC

3. The British Museum Catalogue

4. G.K. Hall reprints of major American libraries (e.g., University of Illinois; Rare Book Division, New York Public; Folger; Berg Collection; Berkeley; Union Theological Seminary)

5. Morgan's Bibliography of British History (1700-1715)

The following ancillary listings were also used:

6. Henry Hills' booklists (as described in my article in Words, 1974)

7. Booklists of Thomas Speed and his successor, Anne Speed (the major publisher of Dawes)

8. The contents of The Whole Works 1732-3 published by J. Wilford

The collocation file was maintained in presumed chronological order of publication, and analysis of the printout made with it revealed that Hills pirated every new sermon by Dawes from 1705 to 1709, but ignored all those before 1705. (He was, of course, prevented by the Copyright Act from pirating those published after 1709). He also ignored his Christian "courtesy books", The Duty of the Closet and The Great Duty of Communication Explained, and his one poem.

The list of publications before 1705 contains 17 items, 8 of which are not in NUC, and 12 of which are not in the British Museum Catalogue (the most extensive collection among the individual libraries listed above). Nine of the items were probably of the 1694-1700 period, but only 5 of them are listed in the second edition of Wing, and none of course in Morgan. The list of publications (other than by Hills) between 1705 and 1709 totals 19, and Hills' piracies total 28. The list of publications from 1710 to 1731 total 24. Coverage by the macroscopic tools is not much greater than for the pre-1705 period. But of particular interest are the reasons for the gaps.

One major reason is that for about a dozen items, no copy has as yet been found in any library. Indeed, it may very well be true that some of them will never be found. However, although the third edition of Dawes' only poetical publication An Anatomy of Atheism escaped even the thorough searches of David Foxon (it does not appear in his bibliography), the copy in The University of Saskatchewan in Saskatoon demonstrates that many will eventually be discovered in lesser-known libraries.

But another major reason is that a catalog entry may often subsume more than one edition of a publication. In NUC, for instance, the first and second

editions of a sermon preached on November 5, 1696 are inadvertently given one entry, even though they are distinguished in at least one of the sources of cataloging data - the published catalog of the Union Theological Seminary Library.

NUC also tends to subsume 4to and 8vo editions of the same sermon simply because contributing libraries supply only the height of the publication in centimetres. The distinction is in fact important, as the collocation file provides evidence that Anne Speed was issuing the individual sermons in aggregated form as "Collected Works". Hills imitated (or perhaps even initiated) the practice.

A source of difficulty for the scholar using the macroscopic tools is the amount of "noise" or the presence of misleading or redundant entries in the lists. The first edition of Wing, for instance, listed 6 Dawes items. An erroneous date in the first item disguised the fact that it was identical with the second. A conjectural date in the third item is certainly a decade too early, and it is interesting to find that the entry was dropped in the second edition. NUC in particular has many multiple entries (three in the Dawes sequence) as a result of careful attempts to reconcile incompatible catalog records from different libraries. It also becomes the victim of contributing libraries which make false attributions (there are two in the Dawes sequence) and make main entries for parts of books (one example in the Dawes sequence).

Another cause for difficulty is the creation of "ghosts" (i.e., entries for which no book has ever existed). NUC contains several in the Dawes sequence. Morgan in particular tends to describe books he has not seen, and although our collocation file contains a dozen unlocated items, the source of our information about their existence is cited. Morgan, however, lists conjectures as facts, especially when dating a particular edition. But the scholar quickly finds out that if Morgan's description of the imprint of a particular book is fairly elaborate, it is good evidence that he had seen a copy of it. He compensates admirably for the cavalier fashion with which most library catalogs treat imprint information and format.

The most important conclusion that we can reach from such close study and comparison of macroscopic tools is that what the scholar needs is not just one all-embracing bibliography with one consistent set of standards of description, but a whole battery of such tools, each with its own (hopefully consistent) standards derived from a clear understanding of the particular purpose to be served by that tool. Union catalogs, library catalogs, enumerative bibliograhies, booksellers' catalogs, publishers' lists, inventories of various kinds - all are grist to the mill.

And even a machine-readable short-title catalog with associated control devices is no substitute for this variety of tools with different functions. But I hope, by giving concrete examples of the use of various methods of providing a scholar with appropriately specialized, limited, macroscopic tools that I have, at least in the case of the biography of a printer and publisher, demonstrated that the theoretical bases upon which the HPB project is developing are both sound and useful.

Lecture 4

BIBLIOGRAPHICAL CONTROL OTHER THAN BY SUBJECT CLASSIFICATION

In order to illustrate the point that subject indexing and subject classification are only two of a number of possible means of bibliographical control of early books, and in many cases not the best possible means for a scholar in the humanities, I propose to expand on the classification by form that seems to have been Henry Hills' "natural" division of his published output. But to demonstrate other kinds of helpful classification or ordering of bibliographical entries, and even some overlooked kinds of bibliographical tools, I propose to choose a more fruitful field than the biography of Henry Hills. I am sorry that I cannot choose a topic with even the tenuous link with India that Henry Hills provides, but I shall assume that, as my audience has an interest in those institutions dedicated to the housing of the documentation over which we wish to have bibliographical control, I may choose one such institution in the 18th century which has not been given proper attention by scholarly researchers. I refer to the coffee-houses of London, England, some of which had quite extensive libraries, and all of which had a profound influence on the social and intellectual life of the time.

But before turning to this unusual topic, I would like to touch on some aspects of Hills' published output that might require very special kinds of classification or ordering in order to provide a scholar with the optimal degree of bibliographical control that the manager of our information system could provide.

I am well aware that if my choice of the biography of Henry Hills as a case-study for illustrating problems of bibliographical control has raised the interest of my audience in Hills himself, you will be disappointed that I am not reporting on the probable results of scholarly investigations into his life and works. But that must await much more work in depth not only on Hills himself, but on all those with whom he was connected during his lifetime, including the authors he published. Once again, it is not our purpose here to report fully on the results of such biographical research. Only so much of the tentative findings will be reported as is necessary to make our points about the provision of the means of bibliographical control for such research.

Ordering of Bibliographical Data for Biography of an Author

Let us imagine that a scholar wishes to write a biography of one of the first preachers whom Hills pirated - Thomas Knaggs. What kinds of bibliographical information would he like to have about Knaggs' life and writings? And in what order of presentation would he like to have it?

First, he would probably like a list of all previous work done on Knaggs. A conventional literature search would produce nothing. The best "biography" (if such it can be called) would be found in Venn's Alumni Cantabrigiensis. If we expand the abbreviations in that entry, we would have the following

account of the man:

> Thomas Knaggs was admitted sizar at Emmanuel College, Cambridge, on June 1st, 1676, having been previously of Trinity College, Oxford. He came from Durham, where he had attended Durham School. He matriculated in 1677, and graduated B.A. in 1679/80. He was granted his M.A. in 1683. In December 1681 he was ordained deacon at York and became vicar of Merrington, Durham, in 1682, and continued to be so until 1720, when he became vicar of Croft, Lincolnshire.
>
> From 1687 to 1697, Knaggs was afternoon lecturer of All Saints' Church, Newcastle-on-Tyne. He was (during an undetermined period) Chaplain to Lord Grey. He was also (for an undetermined period) lecturer at St. Giles-in-the-Fields, London. He died May 12, 1724, aged 63.

Venn's compilation also has an entry for his son, which provides a hint of a tragic personal experience in 1709 for the father, who at that time was presumably 49 years old.

> Thomas, junior, was baptised at All Saints', Newcastle-on-Tyne on March 5, 1690-1. He was admitted pensioner at Emmanuel College, Cambridge, on July 8, 1708, but died in College the following year, and was buried July 31, 1709.

Apart from the results of a literature search for biographical work on Knaggs, our scholar would like a bibliography of Knaggs' publications, and we could oblige him with a short-title catalog listing from the HPB project's files in, say, chronological order.

We would also inform our scholar that special collocation files could be made to increase his bibliographical control of the material. What would he find most helpful for his purpose?

Upon glancing through the printout, our scholar might say that he would like to have the items arranged NOT in chronological order of publication (which may be suitable for studying a printer), BUT in chronological order of composition (which is more appropriate for a sermon-writer). However, there are not sufficient data in STC form to permit this to be done. Nevertheless, our scholar could be informed that if the HPB entries were used as an index to NUC through the HPB/NUC Index, he would find a fairly large number of entries in NUC that contain information about the day on which the sermons (for all the items are sermons) were delivered. By annotating the printout with these dates, a collocation file could be generated. For those items not in NUC the British Museum catalog and other sources could be used. The resulting collocation file could be sorted by machine, but as it is so short, a manual ordering might be all that is necessary in this particular case. The printout generated from this collocation file would reveal that 25 sermons by Knaggs had been published (some in more than one edition) and that dates of delivery ranged from February 5, 1688/9 to June 24, 1722.

Upon studying this new printout, our scholar would probably express the need to have some information about where, and before whom the sermons were delivered. Again, all he need do is annotate the printout. (The annotations could be incorporated in the collocation file if he so wished). He could also note down the special occasions on which the sermons were delivered. The result would be that he would have more precise dates for Knaggs' term as lecturer at St. Giles-in-the-Fields (May 19, 1706-January 30, 1716/17 at least), and he would be able to deduce his whereabouts at various times in his career. Although most of his sermons were delivered to his parishioners, at least five were for special audiences, and the range of occasions that justified his publishing his sermons serve to suggest personal characteristics to his biographer. (Our scholar would no doubt want to use Letsome's 1751 index to complement his bibliographical information at this point).

But what about the subject-matter of his sermons? The subject can frequently be inferred by correlating the occasion, and the text chosen for the sermons. An ordering according to biblical text (as in Letsome) would be of greater use than a conventional subject classification. The subject of a sermon can also (sometimes) be inferred from the title. Only 13 of the 25 sermons have provided sufficient cataloging data for NUC to be useful in this respect, and at this point our scholar must leave the bibliographical framework we have provided for him or her, and must consult the original publications for himself. Moreover, apart from subject-matter, consultation of the originals will reveal biographical information not normally provided in any bibliographical entry except a full-scale descriptive bibliography. An exception to this is NUC K0915205, where the following wording appears in the body of the entry:

> By Tho. Knaggs, M.A. Lecturer of the Parish [i.e., St. Giles], and Chaplain to the R^{t} Hon^{ble} George Lord Bergevenny.

I do not for a moment suggest that such information be provided in the body of a catalog entry (indeed I find that traditional repetition of author information in the title part of a catalog entry is an expensive kind of redundancy). But the scholar may well wish to derive notes from the biographical information on the title page to add to his annotations. He may wish to apply the same treatment to information about the person or persons to whom the printed form of the sermon is dedicated. This is information of particular interest to our scholar; nevertheless, such data should not appear in a catalog whose major functions are to identify and locate a bibliographical unit and relate it to other bibliographical units.

Our scholar's own annotations, however, will have established that Knaggs was chaplain to Lord Grey in 1693; to Ford, Earl of Tankerville from 1695 to 1701 (biographical work will establish that this was the same man); to Fulk, Lord Brook from 1702 to 1710; and to George, Lord Bergavenny from 1714 to 1720. And the St. Giles-in-the-Fields lectureship extended from at least 1706 to 1720. The names of Knaggs' dedicatees will also provide further clues for biographical work on his friends and circle of acquaintance. On the other hand, the wide range of publishers revealed in the

imprints of his published sermons will suggest that he had no close connections with a particular publisher and may also explain why Hills pirated only 6 of the sermons.

Ordering of Bibliographical Data for Source Study

Before we turn to the subject of coffee-house libraries, one more example of scholarly investigation into Hills' activities might serve to show the importance of bibliographical control other than through subject classification. Let us imagine a scholar who wishes to determine the particular source of any one or more of Hills' piracies. For most of the sermons, the matter is simple - only one separately-published item preceded Hills' piracy. And this is also true for many of the poems. HPB listings of such items (or for that matter a listing of Foxon's entries) can be readily made by machine (or in the case of Foxon by manual look-up). But of particular value to the scholar might be a listing of all Hills items for which no similarly titled item can be retrieved. The scholar could then go about the systematic analysis of collected works and miscellanies.

Let us illustrate from both sermons and poems.

John Tillotson (1630-1694) who attained the highest ecclesiastical office before his death, might be a good example for a sermon writer. Wing lists 95 items by Tillotson before 1700, but the HPB files can be used to elucidate them more meaningfully. For it is unfortunately true that one Hills piracy of the early 18th century has slipped into Wing through false dating by the originating library, several items are not listed, and the complicated relationship between individual "bibliographically distinct volumes" and their aggregated form as part of "Collected works" is somewhat confusing in the simple listing. Not all post-1700 editions of Tillotson items have been added to the HPB files as yet, but let us assume that we could oblige our scholar with a comprehensive listing in chronological order up to 1710.

Hills printed 11 of Tillotson's sermons as suggested by his booklists and confirmed by rigorous search for originals (with the help of Letsome). Every one of the sermons had been separately published (several in two or more editions) before 1700. But what was Hills' immediate source? Our scholar, upon analyzing the printout would discover two sets of collected works - The Works 1696 (Wing T1185) in folio, which had run to a 6th edition by 1710; and a 14-volume 8vo edition of the sermons edited by Ralph Barker that appeared volume-by-volume from 1695 to 1704, some of them in more than one edition. Our scholar would be well advised to compile an index to the contents of these two publications in the form of the biblical text upon which the sermons were based. This would enable him to use Letsome to good effect, for the title and even the occasion for delivery is only of secondary value in identifying the sermons. (This index could be made in the form of a collocation file, but on the whole would be just as useful in manually-produced form).

The index would reveal that Hills ignored Barker's painstaking pioneering effort to pass on to posterity Tillotson's oeuvre, and confined his interest to sermons in the folio edition. Our scholar now has to determine which of

the six editions in the HPB printout is likely to be the one used by Hills. He might guess that it would be the 4th or 5th as being contemporaneous with the piracies, but he would have to use techniques of textual analysis on all texts before he could draw the proper conclusions. For instance, Tillotson's Of Evil Speaking is included with a sermon by Sharp in Two Discourses 1698. This pamphlet is not in Wing, but may well have been the source of Hills' piracy of both sermons. To steer his way carefully through the very complicated bibliographical wilderness, our scholar may need the flexibilities provided by our system of collocation files. Once he has determined whether Hills used only an edition of the folio works, he will be able to infer the principles of selection which governed the pirate. If the field of choice was broader, he will have more difficulty but the pinpointing of the time of the choice by correlating the appearance of the first editions of Hills' piracies with the references to them in Hills' own booklists may help make his work easier. The kind of bibliographical control we are able to provide will in fact facilitate his appraisal of the subject-matter of the sermons without actually providing subject access.

To give another example from the poetical pamphlets, our scholar may be supplied with listings of poetical miscellanies that will improve on the macroscopic control provided by Arthur E. Case's bibliography, or the fuller listings provided in The New Cambridge Bibliography Vol. II columns 327-430. As explained in The HPB Project: Phase III, pp. 159-160, the manually-produced records for a first-line index to the contents of poetical miscellanies of the period are already in existence, and could be made machine-readable as soon as time and money become available. This extension of bibliographical control from the macroscopic to one form of the microscopic would provide the scholar with the information that some of Hills' pamphlets are derived from miscellanies, and not from separately-published pamphlets so carefully described in Foxon's book.

As with the sermons, Hills choice of copytext for the poems should provide the scholar with insight into Hills' motives for piracy. His field of choice must be determined first, however. And it might be as well to issue a caveat against jumping to conclusions too fast, even though we may be running into the danger of transgressing the fine line between providing bibliographical support for the scholar, and doing his work for him.

A hint concerning Hills' reasons for expanding his printing activities from sermon reprints into the field of poetical pamphlets can be derived from an advertisement appearing on p. 16 of what may be the two earliest of the verse pamphlets - Nahum Tate's A Congratulatory Poem and the anonymous Windsor-Castle[1]. The advertisement reads:

> Whereas the Printer hereof did receive two Letters by the General Post from an unknown Hand, the last dated July the 1st, 1708. If the Gentleman that sent them shall be pleased to communicate any such Copies as there mentioned, they shall be justly and faithfully Printed and Published, and the Favour most thankfully acknowledged, by
>
> H.H.

---

1. It is also on p. 11 of John Gay's Wine 1709.

An "unknown hand" seems to have offered Hills copy for reprinting, and as the advertisement appears in verse pamphlets, we may conjecture that the unpublished material was "a copy of verses" as contemporary jargon would put it. As Tate's poem celebrates the Battle of Oudenard, fought July 11, 1708 N.S. and Windsor-Castle contains allusions to events in the 1708 campaign no later than July, we may date the beginning of Hills' verse pamphlet printing to August 1708 with some confidence. Because of the nationalistic subject-matter, our scholar might be tempted to hypothesize that his original intention was not to pirate other men's property, but to print and publish "justly and faithfully" any topical verse that came to hand. However, the first three poems in his his first list (the two mentioned above and John Gaynam's Marlborough Still Conquers) were all separately published before Hills reprinted them (Cf. Foxon T55, W525, G93) and all late in July or early in August. Foxon's "neutral" descriptions thus are a very important bibliographical support for the scholar's acts of appraisal.

Other intentions are implied in the first half-dozen publications however, and must not be brushed aside. In the imprint of Windsor-Castle appears the words "For the Benefit of the Poor". This phrase occurs in a small number of later pamphlets too.

Another example is the existence of evidence that Hills, despite his "brown sheets and scurvy letter", took some care in the choice of his copy-texts. A case in point is a claim made on p. 24 of his reprint of Dryden's Absalom and Achitophel:

> To prevent the Publicks being impos'd on; this is to give notice, that the Book lately Publish'd in 4to is very Imperfect and Uncorrect in so much that above Thirty Lines are omitted in several Places, and many gross Errors committed, which pervert the Sence.

Again a selection of items containing booksellers' advertisements would be useful to the scholar.

Our scholar may also have to assume that Hills' motivations may have been both complex and changing.

Foxon's book will prove to be of considerable assistance to the scholar seeking bibliographical control over all possible sources of the text for Hills' poetical pamphlets and, like Wing and Morison, may serve to suggest possibilities for retrieval of data by machine.

Foxon's interest in Henry Hills permitted him to extend the scope of his work specifically to list exhaustively all of Hills' piracies. He normally excludes poems written by authors who died before 1701, but makes an exception for Hills' reprints, for each of which exceptions he notes that the entry is "included here to complete the list of his piracies". Nevertheless, at least one item (Rochester's Poems on Several Occasions) has been overlooked. It was pirated by Hills from Jacob Tonson's 1705 edition[1], the

---

1. See David M. Vieth, Attribution in Restoration Poetry, A Study of Rochester's Poems of 1680 (New Haven and London, Yale University Press, 1963), p. 12 (Cf. Checklist p.501 for bibliographical details).

first edition of which appeared in 1691. Foxon normally includes collections of verse by a single author who died after 1700, but excludes works "only published in miscellanies and in periodicals".

To demonstrate the usefulness and limitations of Foxon's work, I propose to use as a sample of Hills' output all those poetical pamphlets which contain his advertisement lists. The sixteen titles are listed in my article in Words, 1974. Three of these 16 items have no earlier edition listed in Foxon: Browne's The Circus (B527-8 and B527-9); The Commoner (C319); and Poems on the death of...Queen Mary (P690). The first two may, therefore, derive from MS, an unlocated printing in a miscellany, or an as yet overlooked or unlocated separate publication. A 17th century source probably lies behind the Queen Mary poems. Foxon notes that 3 other items were originally published in the 17th century: Beaumont's Bosworth Field, South's Musica Incantans, and Garth's The Dispensary. However, Foxon lists (G21, G22) two 18th century editions of the latter work that preceded Hills' piracy. Frank H. Ellis provides us with the most scholarly edition of Garth's poem[1] and says of Hills' piracy, "This edition prints the 144[0] lines of 1700 and 1703 but with readings from 1699^1 and 1699^2, e.g., at 4, 217 'gentle' (the reading of 1699^1 - 1699^2) instead of 'courtly' (the reading of 1699^3 - 1703). The text, therefore, is a conflation of earlier editions with no independent value." For our scholar's purposes, however, the work of the textual editor goes to show that Hills (though mistaken in judgment) took some care over choosing his copytext, or else used a MS source. Such textual work will need to be done for other poems after Foxon's bibliographical work has been taken into account.

For instance, Foxon lists four of the 16 items in our test sample with only one previous printing: The Libertine's Choice (W101 and W102); Ode on the Incarnation (H364 and H365); The Temple of Fame (B551 and B442); and The Yorkshire Racers (P480 and P481). Like the first three items in Hills' first booklist, verificatory textual analysis will probably establish that the earlier printing is the source of Hills' edition.

Foxon provides two or more possible sources for each of the remaining five items in our test sample, thus giving our scholar considerable bibliographical help in his task of establishing textual relations between the bibliographical entities. The most complicated problems are caused by two poems by Daniel Defoe - A Hymn to the Pillory and The True-Born Englishman. Cooperation between a bibliographer and a textual editor is exemplified in Foxon's careful chronological ordering of the different editions in his catalog, for he was assisted substantially in sorting out those editions by Frank H. Ellis when he was editing the poems for the Yale edition of Poems on Affairs of State. Ellis' use of a draft of Foxon's listings permitted the scholar such bibliographical control that he was able, on the basis of textual analysis, to state categorically that A Hymn to the Pillory was a line-by-line resetting of the first edition of 29 July 1703^2.

---

1. George deF. Lord (General editor) Poems on Affairs of State; Augustan Satirical Verse, 1660-1714, 7 vols, Yale University Press, 1963-75. Garth's poem appears in Vol. VI, edited by Frank H. Ellis.

2. Poems on Affairs of State, (Yale edition), Vol. VI, p. 789.

However, the imperfect listing (copies of possibly six separate editions have not yet been located) and the fact that the poem appeared in several anthologies or miscellanies before Hills' piracy, made it almost impossible to establish the copy-text for The True-Born Englishman. This does not, of course, detract from the value of Foxon's substitute for an STC, but does illustrate how necessary is macroscopic bibliographical control to the scholar.

Foxon's listings go a very long way towards helping a scholar explicate the controversy (or rivalry?) represented by the bibliographical tangle centering on Edward Holdsworth's Muscipula and Thomas Richard's reply (subtitled Hoglandiae descriptio).

Once again, the meticulous chronological ordering of the items in Foxon is of optimal value to the scholar.

To illustrate some of the limitations of Foxon's fine scholarly tool (limitations that are fully justified), we may use the fifth and last item among those in our test sample for which Foxon provides more than one entry before Hills' piracy. This is John Philips' Cyder. Foxon's bibliography, splendid piece of our jigsaw puzzle though it is, must interlock with other pieces (in this case, a bibliography of miscellanies) if we wish to have macroscopic bibliographical control of Hills' sources. As I have already mentioned, a bibliography of poetical miscellanies is in the process of being formed from existing bibliographies as part of the HPB project.

Many years ago, I established by careful textual analysis, that Hills' piracy of Lady Winchelsea's The Spleen and Pomfret's A Prospect of Death, issued as a single pamphlet in 1709, was derived from the 1702 reprint of a redaction of a miscellany that had appeared in many different redactions for over 25 years. The 1702 edition was titled A Collection of Poems: viz. The Temple of Death. I propose to take this volume as an example of how our scholar would use Foxon to help him determine the full range of possible printed sources for Hills' poetical piracies. The feature of Foxon's work that is of greatest use here is his first-line index of poems.

One hundred and one poems are listed in the table of contents of the miscellany. Only eight of them are listed in Foxon's first-line index. Five of the eight listings refer us to editions by Hills: An Essay on Poetry (S388 and S389); The Temple of Death (S390 and S391); The Duel of the Stags (H334, H335, H336); The Spleen and A Prospect of Death (F141). This is probably suggestive evidence that the four pamphlets are derived from the miscellany. The other three of the eight listings refer us to editions by other printers and publishers (A54, P246, P733). However, the last of these references is to A Prospect of Death in separately-printed editions. Foxon does not index the item as part of Hills' edition. On the other hand, the second poem in Hills' pamphlet The Temple of Death is indexed, and as it is also in the 1702 miscellany, it seems probable that it was pirated from there. Foxon's inconsistency in providing first-line indexing even leads him into the rash inference that Hills' text (F141) is a reprint of the separate edition of A Prospect of Death (P733).

standard bibliographies, especially those on the microscopic level of bibliographical control.

Other kinds of index involving use of specialized collocation files in conjunction with the HPB/NUC Index could be made to parts of the printed version of NUC, with automatic machine "look-up" of individual items in existing collocation files, but need I comment further on the possibilities? Rather, let me point out the final characteristic of the HPB/NUC Index which permits a link between the macroscopic and microscopic levels. This is the register of locations, which may be used to supplement the locations already listed in NUC. Once again, those parts of NUC which should be given highest priority for broadening the scope for reporting library holdings could be selected for priority HPB cataloging. Some of these parts might be poorly-represented libraries, some of them might be countries outside the United States and Canada (although Canada could be better represented), and some of them might be subject areas. But the Index would selectively overcome the limitations of present union catalogs.

The usefulness of considering the "jigsaw puzzle" approach to even the development of the HPB/NUC Index, and thus to the indexing of selected portions of NUC and its HPB supplements, is, I believe, an emblem of the usefulness in general of the jigsaw puzzle approach to the development of a system of universal bibliographical control. I hope that in the course of these lectures I have at least provided some evidence of the usefulness of considering early books (that portion of the bibliographical universe printed before 1801) as one giant area of the jigsaw puzzle that is a rational division of the total area for universal bibliography, one that can be further broken down into portions for macroscopic treatment. In particular, I hope that I have at least suggested the possibility that conventional subject classification and indexing (including citation indexes) might be supplemented by a new kind of indexing that will provide coordinating linkages between the two levels through machine-readable files of basic simplicity in recording and storing, but of great subtlety and flexibility in use. I hope, too, that by choosing rather unusual examples of research work in the humanities, I have drawn some attention to the peculiar needs of scholars in the humanities, i.e., those needs not held in common with social scientists and natural scientists. In fact, I hope that what may indeed have seemed an interest in the obscure and peripheral at the beginning of this series of lectures may now seem much more obvious and central for those who are truly concerned with Universal Bibliographical Control.

Foxon carefully notes all the variants in the first edition of Philips' Cyder (including the re-issue in Tonson's The Annual Miscellany, 1708), so that our scholar would fairly certainly be able to demonstrate that Hills' piracy derives from the first edition. But the title poem of Hills' pamphlet is supplemented by three other poems, none of which are listed in Foxon's first-line indexes. All three appear in the 1702 miscellany. From such an observation, the scholar will be led to conclude that he must make a list of all Hills items that contain more than one poem, and proceed to supplement Foxon's first-line index in order to have macroscopic bibliographical control. Foxon's Index of Imprints enables a list of Hills items to be made (although a printout from a machine-readable database would make the task less arduous), but as title pages cannot be trusted, the selection of items with more than one poem would have to be made from copies of the originals. The possibility of recording the results in the form of a machine-readable collocation file would undoubtedly be a relief to the scholar faced by so much note-taking. And the printout could then be annotated with the first lines of the additional poems, eventually to be alphabetized by machine. The fact that so many pamphlets contain "fill-ins" from the 1702 miscellany that were used to supplement a poem from a different copy-text[1] would certainly demand some explanation of Hills' interest in the volume.

Some Conclusions from the Hills Examples

Classification, selection, arrangement of bibliographical entries according to principles other than identification of subject content has been concretely exemplified so far by instances of chronological ordering of composition or oral delivery of sermons, geographical ordering by place of delivery, and ordering by canonical position of a biblical text. Apart from ordering by title and author, which have always been part of a subject classification scheme, we have seen examples of various forms of chronological ordering, of which Foxon's bibliography is a fine example, and the source bibliographies for our machine-readable bibliography of poetical miscellanies is another. The chronological ordering of booksellers' lists also underlines the particular importance of chronology to the researcher in the humanities. Foxon's helpful indexes to his author/title/chronology classification do not include a subject index in any real sense of the term, as pointed out in the previous lecture. As both the examples of research into Hills' work are fundamentally based on an assumed form classification, the general conclusion that classification by form and chronology are more important to scholars than classification by subject seems borne out by the examples. But it may easily be charged that the examples are not really typical. An empirical investigation of what might be termed "typical" is obviously out of the question here[2], so I turn to an example of historical research that is not biased by our attention to a particular printer or publisher.

---

1. Another example is Blackmore's The Kit-Cats (B261 and B262). Hills, in his earlier edition (B259 or B260) included only the title poem. In the later reprint he added two poems from the 1702 miscellany.

2. But I would very much like to see an attempt made elsewhere.

Coffee-House Libraries

I suppose my interest in coffee-house libraries was first aroused in 1966 when I was cataloging David Nichol Smith's library in the National Library of Australia. I found a pamphlet which is in fact the first survey of libraries in English - (William Oldys') A Critical and Historical Account of all the Celebrated Libraries in Foreign Countries, 1739. On the title page a contemporary hand had written "Bank Coffee House" and what seems to be an intended list of names. A literature search a few years later of the subject kind familiar to all research workers turned up only two articles on the subject of coffee-house libraries - by George S. McCue in 1934, and Paul Kaufman in 1969.[1] McCue's article cited contemporaneous allusions to the subject, and Kaufman drew attention to collections of pamphlets whose annotations reveal that they were once part of the reading stock of the coffee-house libraries. A scholar who wishes to find more evidence of the existence and characteristics of such libraries needs to improve on the bibliographical control of primary source materials exhibited by these two men.

His first step, of course, is to verify their sources. Very little can be added to McCue's thorough work, but Kaufman's use of his sources overlooks a great deal of helpful material. He discusses pamphlets from three coffee-houses: one from the Bank Coffee House in Threadneedle Street (the same one from which Oldys' little pamphlet on continental libraries came); two from George's, Temple Bar; and approximately 143 from Tom's Coffee House, Devereux Court.

A preliminary verification of Kaufman's sources for pamphlets from George's has uncovered twelve more. They are part of a collection of 50 quarto volumes in the British Library. An invoice presented by I. Kerslake to the Trustees of the British Museum reveals that "A Collection of English Poetry in 57 vols. 4o £9/9/0 nett" was sold to the British Museum on 11 July 1849. Only 50 of those volumes survive under shelf-marks 11630c 1-4; d 1-8; and e 1-18. The remaining 7 volumes have not been found, but it is possible that the pamphlets they contained have been disposed of or dispersed elsewhere in the British Library. A transcript of annotations on items in the collection would make a remarkable bibliographical tool in itself, as approximately half of the nearly 800 pamphlets contain evidence of their origins in a wide variety of repositories. About 87 of them have sufficient evidence to establish that they were once in Tom's Coffee House. Unfortunately, it is not always possible to identify the coffee-house pamphlets, because some have lost the half-title upon which marks of ownership were possibly recorded, and most of those showing marks of ownership have been severely cropped on the fore-edge and head margins.

Preliminary verification of Kaufman's sources for items from Tom's Coffee-House (they are listed in an arbitrary order in the appendix to his article)

---

1. George S. McCue "Libraries of the London Coffeehouses" Library Quarterly IV (October) 1934, 624-627.

   Paul Kaufman, Libraries and Their Users, London, 1969, Chapter IX (i.e., pp. 115-127).

show many inaccuracies in the listings. The three sources cited are the collection in the British Museum mentioned above (I shall call it the Kerslake collection), another British Museum collection (shelf-mark 1493.c.17) and a collection in Wilmarth S. Lewis' library of Walpole and Walpoliana in Farmington, Connecticut.

The latter two collections derive from the sale at Sotheby's October 31, 1960, when part of the Library of the Writers to the Signet, Edinburgh, was auctioned. Kaufman had seen only one of the two rebound volumes shelved under the number "1493.c.17", and has overlooked the significance of two catalogs in MS that are bound in with the items in that volume. The two volumes contain only 40 of the 73 that were once part of the collection, as is revealed by one MS catalog (a 2-page table of contents). The other MS catalog is probably a leaf from an account book kept by the coffee-shop or a bookdealer, as it is an inventory of some kind. Mr. Lewis' two volumes also contain MS lists on the front end papers which probably represent the original contents within the two bindings. The one labelled "Tracts 2" originally contained 48 items (3 are missing) numbered 1-48 in [19th c?] blue pencil. However, what looks like an 18th century list, numbered 1-56, with prices opposite each number is also present, and may represent the collection before it was bound. Mr. Lewis' second volume of pamphlets has a similar list numbered 1-30 with prices opposite. All except 8 of the items remain within the rather tattered binding.

These MS lists constitute both a device for bibliographical control of Tom's Coffee-House Library after dispersal, and a checklist for further systematic searching. The entries are, as in so many inventories, fairly inaccurate references to titles of pamphlets, but like all inventories and tables of contents, should be categorized as a particularly helpful device of bibliographical control to be used with our "useable substitutes" for universal bibliography.

Let us take first what is probably the oldest of the MS lists, the "Accounts leaf", as I shall call it, in British Museum 1497.c.17. It begins with "Brought over 10.3" and lists 24 items (numbered 18-23,23-40) with prices opposite them. These prices (usually 6^{d}) add up, with the 10s. 3d brought over, to £1.3.4. We may assume, then, that it is the second of two leaves. Six of the 24 items may be identified in the surviving BM collection, and 12 are in the 2 page (19 century?) table of contents for the original collection of 73 items. The fact that they occur <u>in the same order</u> suggests that the "Accounts leaf" may have been a sale catalog.

Using the "Accounts Leaf" as a search tool, I discovered in the Bodleian Library, Oxford, a very interesting collection of pamphlets (Firth c.8) arranged in presumed chronological order. An undated item (it is actually 1732) is followed by 69 other items ranging from 1701 to 1760. And among them are certainly 5, possibly 10, of the items in the "accounts leaf". The lack of certainty about Tom's Coffee-House ownership is due to the fact that the dealer who sold the items (presumably to Sir Charles Firth) had very carefully washed away with soap and water as many ink marks as possible before combining the pamphlets with others. One can still feel the soap on one of the pages! Most of the detectable items from Tom's show a speckled red fore-edge and similar marking along the top edge, so the

extensive cropping of top and side edges was probably done before that dealer disbound them. Indeed, Nos. 18, 28, 45, and 57 are the only uncropped items, so the original binding may have contained some items not from Tom's. The Bodleian Library stamp is dated 13 July 1937, so the scholar needs to find out both when and where Firth obtained his collection before that date.

The "Accounts Leaf" used as a checklist in combination with Foxon's bibliography also reveals another depository of Tom's Coffee-House items - the Brotherton Library at the University of Leeds. (There are in fact seven pamphlets there.) This use of two different kinds of bibliographical control device, each compiled for a different purpose, in order to locate an item for yet another purpose, is indeed a fine example for the main theme of these lectures. Foxon notes the fact that the Brotherton Library copy of Macnamara Morgan's The processionade, 1745 (Foxon M457) was once in Tom's Coffee-House only because it provides a date (Feb. 12). His purpose is simply to establish the probable date of publication. If no date survives on a Coffee-House pamphlet, he did not note down the ownership. An amusing (but also frustrating!) note on The Seventh Satire of Juvenal Imitated, 1745 (Foxon S355) records that he had seen Tom's copy with date 31 January - but he has obviously forgotten where!

The other MS lists - the various Tables of Contents - may also be used as checklists, but I would suggest that their main use is not just for tracing down individual items, but for identifying examples of yet another bibliographical control device - the antiquarian booksellers' catalog. For the history of the dispersal of Tom's Coffee-House pamphlets will eventually be traced through such neglected but useful bibliographical tools. They are neglected simply because our bibliographical control over them is very poor, and because their contents are frequently difficult of analysis. Sotheby's sale catalog of the material from the Library of the Writers of the Signet is a case in point. It is very difficult to analyze precisely what was sold from that library. The portions of the collection bought by the British Museum and Wilmarth S. Lewis still leave many questions as to what else was sold, and to whom. Some of the buyers must have been bookdealers, and one would expect to find post-1960 bookdealers' catalogs with some of the items in them. It is too early in my own investigations to reconstruct even in broad outlines the history of the dispersal of the original coffee-house library, but let me illustrate the problem of analysis of bookdealers' catalogs by reference to one of 1968, eight years after the Sotheby sale.

Thomas Crowe, 77 St. Giles St., Norwich, England, issued in 1968 a catalog with an interesting title page reading thus: "A number of pamphlets: many scandalous, seditious, unlicensed" Autumn MCMLXVIII. Norwich. Printed by W.I.J. Mackay, for Thomas Crowe, Bookseller, St. Giles". A note on p. 1 of the catalog tells us "Many of these pamphlets have been circulated from Tom's Coffee-House in Covent Garden (sic!)...". Apart from much misinformed speculation about the coffee-house (it was in fact in Devereux Court, near Temple Bar), the note arouses curiosity as to which of the pamphlets in the catalog were in fact Tom's, and what was the provenance of them. The University of Kansas acquired two of them, both accounts of trials (Sir William Morris, and Sir Chaloner Ogle), both of them with clear evidence that they were once in Tom's Coffee-House. One way to determine which of

dozens of possibilities in Crowe's catalog were in fact Tom's is to attempt to trace the buyers of the pamphlets. If a bookseller keeps records, he must sometimes refuse to divulge information in order to respect his clients' privacy. The scholar must, therefore, use his battery of bibliographical tools in ingenious ways in order to find which library catalog needs to be searched.

For instance, The Table of Contents MS in BM1497c.17 (which I shall call the "Signet Contents" MS) contains 14 items not in BM 1497.c.17. No. 61 (Placebook for the year 1745) may perhaps be Item No. 32 in Thomas Crowe's catalog. Foxon lists the poem as W497, and gives a number of locations, each of which must be checked in the hope that Tom's copy will be found. The search may reveal the location of others of the 14 items in the "Signet Contents" MS, or may identify other Tom's items in Crowe's catalog. To give another example, item 4 in one of the MS lists in Mr. Lewis' collection (In Laudem Eduardi Vernon) is listed as Foxon U26. Foxon had not seen a copy, but found it listed in Pickering & Chatto's Cat 219/7228a. NUC: Pre-1956 Imprints locates another copy in the New York Public Library.

Such systematic work may not always accomplish our purposes, however. It is then that the scholar may have to fall back on serendipity through lack of the proper plethora of bibliographical tools. And we must not neglect serendipity as part of the whole spectrum of bibliographical control devices!

Serendipity has always been a very important element in the research methods of scholars in the humanities. It can be worked up into an art by intelligent use of intuition based on a background of experience of how documents are transmitted over time and space. If we are to find copies of pamphlets that once were part of a coffee-house library, we must look at collections of pamphlets of the appropriate era. Because separate pamphlets are something of a nuisance in a library because of their size, it is very common to find them bound up as "tracts" in thick volumes, and sold as such by booksellers without particularization of title. Indeed, many a valuable collection still needs to have the individual items cataloged. A case in point is the nine linear feet of folios, quartos, and octavos of poetical pamphlets containing 1089 items, the bulk of which were acquired by the University of Cincinnati on 21 April 1933 from Dobell, and supplemented over the next decade or so by small purchases from a variety of booksellers. The collection is entered in the general catalog under "Anonymous" and a shelf-list serves as the major bibliographical control device. However, Foxon's bibliography makes the collection highly available to all those scholars who may want to have access to it. In 1963, Beth Smilay Newman wrote an M.A. thesis "Eighteenth Century Anonymous Political Verses and the Career of Robert Walpole" using 104 of the poems in the collection, but many other theses could be based upon the collections.

A search through the folios and quartos in this collection resulted in finding nine items that may have once been in George's Coffee-House, although none have the characteristic "This belongs to..." formula. The method of numbering pamphlets is the only evidence. The 9 pamphlets date from 1740 to 1754. Serendipitous searching in other collections of a similar kind has so far added an item from the Bridge Coffee House (1749),

five items from Dod's Coffee House (1754-1758), an item from Mamford's Coffee House (1776), as well as additional items from Tom's and George's.

Some conclusions can, I think, be drawn from my example of research into coffee-shop libraries. First, there are in existence a great variety of bibliographical listings and inventories of the kind that would automatically be excluded from the scope of "true" bibliography by such systematic bibliographers as Besterman. Nevertheless, such lists, used in conjunction with the macroscopic "true" bibliographies that we have discussed earlier on, enormously increase our bibliographical control in particular research areas and must not be ignored in the total system of universal bibliographical control. Second, ordering and classification in such tools (sales lists, whether printed or MS, booksellers' or auction catalogs, accounting documents or inventories) is not of the same kind as should be found in tools made specifically as instruments of bibliographical control, so they need very careful analysis before their proper use can be determined. Third, the bibliographical control of such tools has been much neglected, partly because of their lack of convenient arrangement. A bibliography of bibliographies seems called for.

But the coffee-house library investigation also draws attention to the particular emphasis on books as <u>objects</u> (not as the vehicle of <u>subjects</u>) which occupies the attention of scholars. For that reason, you may still believe that our choice of topic for research work is not typical of all historical research. In the history of culture, libraries, and printing or publishing, it is perhaps the greater part of all scholarship, and should not be neglected in designing any grand scheme for universal bibliographical control. Within such a category of scholarship will be found those scholars who require microscopic bibliographical control, which is the subject of my next and last lecture, and this, too, should not be neglected in putting together our giant jigsaw puzzle.

Lecture 5

BIBLIOGRAPHICAL CONTROL ON THE MICROSCOPIC LEVEL

I shall return to Henry Hills in this final lecture in illustration of the scholar's needs for bibliographical information of a kind that only a few Special Collections librarians attempt to provide. Most library collections are content to have only one copy of a particular book, presumably on the grounds of optimization of the use of acquisition funds. But this policy in a Special Collection puts a serious barrier between some kinds of scholar and the full utilization of a collection. As I pointed out in my introductory paper on Hills in Words 1974 (p. 73) a rare opportunity (at Yale University) to compare side-by-side two seemingly identical copies of one of Hills' pirated sermons revealed that Hills was in fact concerned enough about aesthetic appearances that he made several changes in the typesetting, and corrected what may have been an error. At the time of reporting, 28 poetical pamphlets and 15 sermons which seemed to be identical from normal cataloging information had been proved to exist in more than one state, issue, or edition. The work of side-by-side comparison must await an opportunity for the scholar to acquire for himself a control copy, or until a library buys multiple copies.

To illustrate what may eventually result from such comparative work, we could cite several cases in Foxon of Hills' poetical piracies. The Henry E. Huntington Library in San Marino, California, has two copies of Hills' piracy of Sir Robert Howard's The Duel of the Stags 1709. One of them has MS corrections in a contemporaneous hand, and an early cataloger, realizing from side-by-side comparison that they were two different editions, noted on the catalog card: "Different edition, completely reset with some of the MS corrections noted in that copy corrected here". This was a false conclusion based on insufficient evidence. The truth of the matter is that the inferior text (with MS corrections) is a later reprint, not necessarily of the other edition. For Foxon distinguishes two impressions or editions (it is not clear which from his description) of the earlier text (H334 and H335) from the later one (H336). He was unable, from the bibliographical evidence, to determine the priority of the two earlier texts, but the distinction between them was made possible by side-by-side comparison of three copies in the Bodleian Library. His technique of noting the position of the signature on each page undoubtedly helped also, but one suspects that not all the bibliographical work has been completed, for the Bodleian copy of H334 has A3 missigned "A2" (other copies have only A3 signed). H335 has A2 and A4 signed also.

I am sure that the sermons will also reveal many similar problems of identification of editions, issues, and states. In preliminary work on Offspring Blackall, Bishop of Exeter, I have established that at least two undated editions of The Divine Institute of Magistracy can be identified, as well as one dated 1709. The sermon was delivered before the Queen on Tuesday, March 8, 1708/9, and Will Rogers' legitimate editions (one of 32 pages, one of 24 pages) are dated 1709. In order to establish the order in which these editions were published, the scholar will have to supplement a search

for dated advertisements with textual analysis of the relationships between those editions. But he should not carry out textual collation until all possible editions have been identified. Experience teaches the scholar that premature textual collation frequently suggests that missing or disguised editions remain to be found, and that his choice of exemplar for collation is possibly a wrong one. Only after chronological classification of all editions, issues, and states of a particular text have been worked out is the labor of textual collation worth doing.

I own three copies of Blackall's The Subjects Duty, 1709. Side-by-side bibliographical comparison reveals that they are in fact two distinct editions, one of which exists in two states. How many more copies must I compare with my copies before I can be sure that all editions, issues, and states have been identified? Some external help is to be found in Hills' advertisement lists of sermons, after they have been tabulated in chronological order. For we may conclude that Hills pirated 11 sermons and 8 Boyle lectures delivered by Blackall, and that fluctuations in the number available at any one time can be accounted for by new editions being printed to replace sold-out stock. But, eventually, the identification of those editions must be arrived at by careful comparison of type-setting (a bibliographical task) and identification and analysis of textual variants (a textual task). I hope that these examples of problems of analytical bibliography will be sufficient to indicate to you the requirements for multiple copies of a particular book that scholars frequently experience, and that you will accept my assurance that there are many other circumstances in which the scholar will require very precise bibliographical details about many copies of the same book. I shall illustrate his difficulties in using conventional catalog entries, and even descriptive bibliography entries by dealing with what I once called the high point in Hills' printing career - the splendidly-produced edition of Aesop's Fables dated 1687 on the title page.

Hills Polyglot Aesop

Before a scholar can come to any conclusions about Hills' role in producing Francis Barlow's splendid book, he must first find out, from among the bewildering variety of copies, which ones represent the intended "ideal copy". (The term "ideal copy" is a neo-Platonic concept introduced into the field of analytical bibliography in the 20th century to solve a classificatory problem. Does one describe the book in front of one, or describe what the printer or publisher intended it to be, and note actual variations from this ideal that are exhibited in the copy being described?) What kind of bibliographical control will our scholar need to exercise if he wishes to have an example of an "ideal copy"?

In an unpublished paper presented at the Seventh Annual Meeting of the American Society for Eighteenth Century Studies in April 1976, I reported on an unfinished search for sufficient copies of the polyglot Aesops of 1666, 1687, 1703 (and possibly an undated one) to provide a scholar with sufficient information for him to determine what precisely Hills actually printed. At this point in time, I am still not sure that I have examined sufficient copies to provide that information, but I shall give first of all a tentative discursive explanation of the facts so far discovered in order to provide

a background for exploring our present topic - microscopic bibliographical control. This tentative account of the publishing history of parts of Francis Barlow's polyglot Aesop should be taken to be an hypothesis for verification or modification as more exemplars of the text (or texts) of the work become available.

By 1665, Francis Barlow had engraved

1. A plate depicting Aesop among his animals, with a 4-line caption:

   See here how Natures Book unclasped lies,
   Whose pages Aesop reads with pearcing eyes,
   Who wth wise Apologues from Beasts deriv'd,
   Tells man they for his conduct were con-
   triv'd.

2. A plate from which an engraved title page dated 1665 could be printed

3. 110 plates, each an illustration for the 110 fables, each with a set of English verses beneath them (presumably by Tho. Philipott)

Barlow (or a collaborator) had also arranged with Ann Seile and Edward Powell to have William Godbid set type for:

4. 10 folio sheets containing an English life of Aesop

5. 8 folio sheets containing a French life of Aesop

6. $4\frac{1}{4}$ folio sheets containing a Latin life; 3/4 + 54 + 1/4 sheets containing the 110 fables; 1/2 a sheet containing a table of contents.

7. 1 folio leaf containing a title page dated 1666

8. 1 folio sheet containing a dedication to Sir Francis Pruijan and an address "To the Reader"

The sheets containing the fables were designed to have on the recto of each leaf an impression from Barlow's plate (No. 3 above) with a Latin letterpress version (by Robert Codrington, presumably) below the engraved English version; and, on the verso, the French version of the next fable. (Thus, in any particular opening of the book, the French version of a particular fable would appear on the left-hand page, opposite the engraved English version and letterpress Latin version on the right-hand page).

At some time (or times) before 1687, Barlow (or his agent?) probably engaged Christopher Wilkinson, Thomas Fox, and/or Henry Fairthorne to have Henry Hills set type for

9. 10 folio sheets containing a new French life (Cf. No. 5 above)

10. $4\frac{1}{4}$ folio sheets containing a reprint of the Latin life; 3/4 + 54 + 1/4 sheets containing new versions of the 110 fables; 1/2 sheet for a new table of contents (Cf. No. 6 above)

11. 1 folio leaf containing a title page dated 1687 (Cf. No. 7 above)

12. 1 folio sheet containing a dedication to the Earl of Devonshire, and an "Address to the Reader" (Cf. No. 8 above)

In the meantime, Thomas Dudley, with the assistance of Francis Barlow, had, in the years 1678 and 1679 etched:

13. 31 plates to illustrate the life of Aesop (possibly the English, possibly the French?)

Barlow had also reengraved:

14. The wording within the oval centre of his 1665 engraved title page, thus eliminating the date (Cf. No. 2 above)

He (or some other engraver?) also reengraved:

15. The English words under the 110 illustrations of the original plates (Cf. No. 3 above)

Thus substituting verses by Aphra Behn for the original ones by Thomas Philipott. And an unknown engraver etched:

16. A plate containing the Arms of the Earl of Devonshire

As with all descriptions of complex processes, this description, simplified though it is, of the creation of 16 sets of physical objects (or series of objects) from which printed pages could be made, may seem too complicated for oral presentation. To help reduce the information overload, let me further simplify the 16 categories. Seven of the sets of objects are engraved plates (1-3, 13-16); nine are settings of type (Nos. 4-12). Only two of these (Nos. 14 & 15) are a reworking of the same physical object (Nos. 2 & 3), but some changes in the type-setting were made in others during the process of printing the sheets, either in 1665 or later.

Sometime in the late 1680's an unknown number of copies of sheets were printed from items 9-16, which, together with surviving copies of sheets printed from items 1-8, were used to produce a new issue of the 1665 book. It was not strictly a new edition, for many of the sheets were from the original edition. In particular, the sheets of the English Life were incorporated in all copies of the new issue, but some sheets from the original edition of the fables themselves were issued in some copies of the new issue. Some of the copies were on large paper, some on small or common paper, as in the original edition. Study of some surviving exemplars of the text reveal that some of the sheets (both original and reprinted) exist in variant states. And in most copies of the small paper version, many sheets are survivors from Godbid's printing.

All copies so far seen of the large paper version with the 1687 title page contain Hills' reprintings (Nos. 9-12 above) rather than Godbid's (Nos. 5-8 above). Nevertheless, variations have been found in the order in which the sheets appear within the binding. However, classification of variants in the small paper version is complicated by the existence of the re-impressions of Barlow's plates in a variety of ways, together with re-imposition of some of the letterpress in the preliminary sheets. (I apologize for the complexity of the information above, but it has in fact been greatly simplified by using an implicit classificatory system based on the physical process by means of which the surviving exemplars were produced.)

But we are not finished yet. Three copies of a 1703 re-issue so far examined suggests that Richard Newcomb merely set type for:

17. A new title page dated 1703

18. A reprinting of No. 12 above (the dedicatory sheet)

and reissued the remainders of the small-paper issue in all its complexities, including sheets surviving from 1666.

This tentative discursive account of the history of the polyglot Aesop in roughly chronological order has been made possible by a classification of variations discovered in a fairly large, but not yet large enough, number of copies of a book that will appear in conventional library catalogs as three "editions" of 1666, 1687, and 1703. This does not mean that I am asking librarians to provide further details in their catalog cards for use by scholars. I wish to emphasize that it is by proper provision of a number of bibliographical tools, each with a different function, that the scholar is able to exercise partial microscopic bibliographical control over a very small portion of the bibliographical universe. In this particular case, both the inadequacies and the special value of three different kinds of "device for bibliographical control" can be exemplified.

Bibliographical Control of the Polyglot Aesop

The three kinds of aid have been provided by librarians, bibliographers, and booksellers, and each in their own way contribute to improving our bibliographical control over the first polyglot Aesop.

Consultation of The National Union Catalog: Pre-1956 Imprints provides us with a reproduction of a New York Public Library catalog card for the 1666 edition (dated November 29, 1948) and a Library of Congress card for the 1687 edition (apparently dating from 1918). Both give much more information than one would normally expect. The 1687 card even includes a description of the 1666 title page (with, it must be said, rather unsettling differences in detail about the imprint). The statement that the first edition was largely destroyed in the fire of London does not seem to be an adequate explanation for the peculiarities of the "second" edition, but the apocryphal story provides food for thought. This attempt to elucidate the bibliographical relationship is traced to the Huth Library Catalog, so the librarian has taken note of the work of the bookseller who published that catalog.

Indeed, the sale catalog for the Huth Library (1911) gives us much more accurate information about imprints than is to be found in the library catalog cards, but the NYP card is better than the LC card in this respect. Information about the binding of each copy is given with approximately equal accuracy and detail in our examples of both conventional library catalogs and a sale catalog of a private collection. Transcription of title page wording is slightly fuller in the Huth catalog, but essentially agrees with the library cards. The information in the catalog cards (31cm and 32cm) for the height of the books may permit us to infer that they are Small Paper copies, for Large Paper copies tend to be about 34.5cm. The Huth catalog not only identifies the format (folio) but also states that the 1687 copy is Large Paper. Collation seems to be of about equal quality for the period at which the cataloging was done, but library cataloging gives more attention to pagination, sale cataloging to signatures. The Library of Congress card attempts to conflate the two systems of reference, thus providing more valuable information to the scholar who does not have a copy of the 2 editions before him.

Finally, the auctioneer's "blurb", which decorously attempts to augment the value of lots 64 and 65, on the first day of the sale, provides at least some bibliographical data on the relationship between the two "editions" which a sceptical scholar may use for a working hypothesis.

The library card for the 1666 edition also contains references to standard bibliographical descriptions (Brunet, Grolier Club's Wither to Prior catalog) and to a history of engraving - Sparrow's British Etching. Such helpfulness is much rarer in modern library cataloging.

The Grolier Club publication is a model of the kind of thorough bibliographical description being developed in the first decade of the 20th century. It is a prototype of the kind of publication which the Greg-Bowers tradition has been attempting to elaborate since the foundation of the Bibliographical Society in 1892. Thus, it contains more elaborate descriptions and commentaries than are to be found in early 20th century library cataloging or in contemporary booksellers' catalogs. The collations, surprisingly, omit all reference to pagination, and much of the descriptive material is impressionistically derived in an unsystematic fashion from close study of the book. The note that Plate 17 (among the 31 illustrations to the Life) is usually wanting, "being generally removed as of too free a

character", explains why the Huth catalog stresses its presence in the copy being offered for sale. An interesting note is also added that Barlow's plates were used (with 4 of the 31 illustrations of the Life omitted) in a 1714 French edition at Amsterdam "Aux depens d'Etienne Roger" - an edition listed in NUC (A0089091).

Brunet, in contrast to the Grolier Club elaborateness, identifies with admirable economy the books under discussion (reminiscent of a short-title catalog entry) but does not attempt to describe them. Instead, his entry is designed to guide the bookcollector on what he should pay for the books, and why. The commentary therefore provides curt but useful comments, which not surprisingly, are more specific (and slightly more accurate) than those derived from them in the LC card. The new material in the 1687 edition is much more circumstantially and accurately defined than in either the Huth entry or the Library of Congress entry.

As can be seen, the primary function of the library catalog card (to guide a reader to the book in a particular library, and to relate it to other works in the library collection) has been extended to incorporate some information (not always successfully) from other sources of bibliographical data. It is in fact modern practice not to add much in the way of summary, and this is wise. It is much better to provide accurate references to those sources so that the user who wishes to, may look them up. The bookseller, of course, provides accurate description of his wares to avoid any possible return of books bought under misapprehensions, but he is also diligent in finding out special significance for each of his items, and thus, provides stimulation for scholarly elaboration or rebuttal. The bibliographer of the Brunet tradition merely provides an extension of the work. And the descriptive bibliographer elaborates descriptions to the point where the analytical bibliographer is able to optimize his microscopic bibliographical control.

It is not the place here to provide a sample of a bibliographical description of the 1666-1687-1703 work with its various texts and exemplars of the text, especially as not enough exemplars have been examined, but perhaps a tabulation of the 1666 sheets found in the 1687 and 1703 issues of the text of the fables themselves will provide something of an indication of the complexity. All four British Library copies of the Large Paper impression of the 1687 issue (including one owned by the Earl of Devonshire, to whom the issue was dedicated) and two copies of the Small Paper issue in the Henry E. Huntington Library and the Library of Congress contain no 1666 sheets in the Latin life and text of the fables themselves. The other copies so far examined were found in five libraries:

The Chapin Library, Williams College, Williamstown, Mass.

New York Public Library, New York City

The British Library, London, England

The Folger Shakespeare Library, Washington, D.C.

Henry E. Huntington Library, San Marino, California

This simplified table does not do justice to the many variations between

1666 Sheets in 1687 and 1703 Issues of the Polyglot Aesop

	B	C	D	E	F	G	H	I	K	L	M	N	O	P	Q	R	S	T	V	X	Y	Z
1687 Chapin					X	X	X	X			X					X		X				
New York Public					X	X	X	X			X					X		X		X		
British Library					X	X	X	X			X					X		X		X		
Folger						X	X	X		X											X	X
1703 British Library					X	X	X	X			X	X							X	X		
Folger	X	X	X	X	X	X	X	X	X		X							X		X		
Huntington					X	X	X	X											X			

	Aa	Bb	Cc	Dd	Ee	Ff	Gg	Hh	Ii	Kk	Ll	Mm	Nn	Oo	Pp	Qq	Rr	Ss
1687 Chapin				X	X	X	X				X							
New York Public				X	X	X	X				X							
British Library				X	X	X	X	X										
Folger			X	X	X	X	X				X							
1703 British Library		X		X		X	X				X		X					
Folger				X		X	X				X		X					
Huntington						X	X	X			X						X	

	Tt	Vv	Xx	Yy	Zz	Aaa	Bbb	Ccc	Ddd	Eee	Fff
1687 Chapin						X	X	X			
New York Public				X		X					
British Library					X						
Folger											
1703 British Library											
Folger								X			
Huntington											

	Ggg	Hhh	Iii	Kkk	Lll	Mmm	Nnn	Ooo	Ppp
1687 Chapin									
New York Public									
British Library									
Folger									
1703 British Library									X
Folger									
Huntington		X		X					

copies that would be of use to the scholar seeking to elucidate Hills' role in the publication of the polyglot Aesop. The Henry E. Huntington Library copies of the 1687 and 1703 issues may be taken as an example. The 1687 "edition" (despite the fact that all the sheets seem to be from the same setting of type and same plates as in the large paper copies) seems likely to be a "made-up" copy put together, possibly in the 19th century, from two possibly defective copies. For instance, Plate 17 (the one frequently missing from other copies because of its alleged pornographic character) has been replaced by an extra copy of Plate 23. Again, the 1703 issue contains a defective impression of Hills' sheet R, with blank space where the plates should be. A copy of Godbid's sheet R has been used to repair the omission. The whole of the first leaf has been bound in, but only the illustration (cut from the second leaf) has been pasted into the blank space on the recto of R2. For that matter, extra plates (possibly from the 1666 impressions) have been pasted onto the verso of kkk2, obscuring Hills' text, and bound in between sheets Ooo and Ppp, thus duplicating the illustration on Pppl.

Tentative Conclusions from an Unfinished Search

Conclusions can be drawn from examination of the dozen or so copies reported on above, but even the patterns in the table should be sufficient to demonstrate that more copies must be found before such conclusions can be trusted. The scholar, therefore, needs bibliographical tools that will help him locate further copies. Those already mentioned above as descriptive tools also serve as location tools.

NUC may seem to imply that the New York Public Library copy of the 1666 edition is unique in North America, but Wing's STC lists copies in the Chapin Library, the Folger, and the Huntington. The Chapin copy is in fact a large paper copy, the others a small paper copy. Wing also lists copies in the British Library; Oxford and Cambridge Universities; Trinity College, Dublin; and the Bibliothèque Nationale, Paris. We cannot tell whether they are large paper or small paper copies.

NUC lists seven copies of the 1687 issue in the United States; Wing of course limits himself to five on either side of the Atlantic. Wing also lists an undated "edition" in the Huntington Library which, in fact, is a copy of the 1703 issue with the date erased from the title page.

The present whereabouts of the Grolier and Huth copies would probably be traced by a dedicated scholar, but it seems fairly certain that many other copies will also be found by systematic searching. The need for comprehensive location information (and not Wing's selective listings) for any STC project should be demonstrated by this example, and the sophistication of union catalogs in machine-readable form will hopefully provide such microscopic bibliographical control in the future.

At the present stage of development of bibliographical control over the polyglot Aesop, we may conclude that if a scholar wishes to carry out side-by-side comparison of a large paper and a small paper copy of the 1666 edition, we know of only one library where it can be done - the Henry E. Huntington Library. The comparison of these copies reveals minor differences

in type-setting in the inner formes of gatherings B and C and the outer forme of B of the English life. The small-paper copy has the corrected and, therefore, final version of the text. All other copies so far seen seem to confirm this conclusion.

Other variations are to be found, too. For instance, in sheet Gg of the polyglot text, the illustrations for fables 48 and 49 have been impressed in the wrong order in the large-paper copy, and corrected in the small-paper copy. Similarly, in sheet Ss, the illustrations for fables 70 and 71, in the wrong order in the large-paper copy, have been corrected in the small-paper copy. The confusion in the latter case has been increased by an erroneous running title at the top of p. 143 (recto of Ss2) inadvertently left in the skeleton when the inner forme of sheet Rr was completed. Such conclusions of analytic bibliography which can be carried out on one copy of a book are not, however, relevant to our present purpose, which is to demonstrate the necessity for multiple copies in the work of microscopic bibliographical control. Nevertheless, comparison of copies is often the fastest way for a scholar to detect the anomalies in a particular copy that he might discover by applying analytic methods. An example can be given from comparison of a copy of the large-paper impression of the 1666 issue of the polyglot Aesop, and a large-paper copy of the 1687 one.

Unfortunately, no library has yet been identified that contains a large paper copy of both issues, but as comparison of large-paper and small-paper copies of the 1687 issue in the British Library suggests, the small-paper copy in the Huntington and the Library of Congress is probably from the same setting of type and plates as large-paper copies. Given the lack of certainty because side-by-side comparison has not been made, we may assume that a comparison between the large-paper 1666 issue and the small-paper 1687 issue at the Huntington is the next best thing to what we need to do. The comparison reveals that Hills (or an editor giving him directions) noticed that the illustration for fable 16 in the 1666 edition was not appropriate for the French and Latin texts; that the French text of fable 68 does not fit the illustration (including its English text) or the Latin text; and that fable 104 seems to be a variation of the texts of fable 16 and the French text of fable 68. He solved the problem by transferring the illustration of fable 104 to fable 16 in the 1687 reprint, by providing a radically different French text for fable 68, and by using a completely new fable 104 to elucidate the illustration wrongly used for fable 16 in the earlier printing, and now applied to 104.

Comparison also shows that Hills made necessary corrections while reprinting, but also perpetuated errors (such as numbering fable 5 as VI). He also introduced new errors, which we need not go into here.

Comparison of many copies of the 1687 and 1703 issues will also be necessary if the analytical bibliographer wishes to clarify, modify, or expand the tentative account of the publishing history of parts of Francis Barlow's polyglot Aesop with which we began this portion of the lecture. You have probably had enough detail already, but one last example should clinch my argument. Comparison of copies of the 1687 and 1703 issues suggest that an even more complicated history of the dedicatory sheet (Nos. 8, 12, and 18 above) will eventually emerge from comparison of other copies.

Some copies of the 1687 issue include a text of "To the Reader" signed "Francis Barlow" on the recto of the second leaf of the sheet. The verso is blank. Other copies with a blank verso have an unsigned text on the recto. Yet others with a blank verso have an unsigned text with completely different "accidentals" such as spelling and punctuation. Others have one or other of the unsigned texts, but with an impression on the verso from the plate depicting Aesop with his animals (No. 1 above). No copy has yet been found with both the separately-printed leaf and with the last of these variants. The 1703 issue usually does not contain the separate leaf. No plausible explanation of these variants is possible after so few copies have been examined. I will not supply the information in tabular form, as by now you should have ample evidence from this example of the need for microscopic bibliographical control of small portions of our bibliographical universe.

Relationship Between Macroscopic and Microscopic Bibliographical Tools

Response to the need of scholars for descriptive and analytic bibliographical tools is usually provided by scholars rather than librarians, and by descriptive and analytical bibliographers rather than general humanist scholars. But even though during the past four or five decades analytical bibliographers have devoted their attention to devising precise and systematic methods of recording bibliographical information for highly specialized use by textual editors and other scholars, librarians must not abdicate their responsibility to link their work up with that of the bibliographers. Librarians are increasingly becoming responsible for creating the major "bibliographical instruments" that will increase on the macroscopic level bibliographic control of early books, but such noble efforts as the creation of national bibliographies, union catalogs, international subject bibliographies, etc. should not be undertaken without a full understanding of the infrastructure of macroscopic bibliographical control which I have tried to suggest by concrete examples, and at least some awareness of the more specialized instruments for microscopic bibliographical control. If a complete system of bibliographical tools (our jigsaw puzzle of universal bibliography) is to emerge in the computer age, the linkages between the macroscopic and microscopic levels must be carefully thought out and made explicit in such a way that the scholar of the humanities as much as the natural or social scientist is well served. I think that it is in this context of universal bibliographical control that our very tentative but nevertheless purposeful efforts in the HPB project should be understood. And for my final attempt to relate the project to universal bibliographical control and to international or global approaches to it, I shall try to explain how the HPB/NUC Index, the general control device among our multiplicity of collocation files for special purposes, can be made to serve as a linking system between the macroscopic and the microscopic.

The HPB/NUC Index

The National Union Catalog: Pre 1956 Imprints is undoubtedly the most comprehensive macroscopic tool for the bibliographical control of early books ever produced. Thus, any machine-based system for macroscopic control should make full use of, and attempt to supplement, improve, or correct it if at all possible. This means that the main file of the HPB project should

not be a substitute for NUC; it also means that the retrieval scheme for the HPB project should, if at all possible, provide means of overcoming the decided limitations of the alphabetical arrangement of NUC. The "fragmentation of thought" that Wroth attributes to man's invention of such a convenient classification system for bibliographical data as the alphabet is certainly obvious when one wants to use NUC for research purposes other than the identification and location of exemplars of texts by a particular author.

For these reasons, the STC surrogate for a bibliographically distinct volume which is the basic element of the HPB main file may be considered to be an index term either to an NUC entry, or to a more elaborate descriptive bibliography entry in a tool for microscopic bibliographical control; or it can be a temporary substitute for a non-existent NUC entry or a non-existent entry in a descriptive bibliography. It should contain sufficient bibliographical data to enable the most-needed forms of arrangement to be carried out by the computer. (And need I say now that chronological order is more important than subject classification?) But an HPB project entry should not be considered to be a substitute for library cataloging, bookseller's cataloging, or bibliographical cataloging. It is short-title cataloging with all the advantages of an identification device (speed of compilation, accuracy within the limits of brevity, economy of data-gathering and storage, simplicity and versatility for retrieval). It also has the disadvantages of an identifying device (it is no substitute for the original book, nor for elaborate descriptions of the book). Because of the entry's brevity and accuracy, comparison of many entries juxtaposed to facilitate the appraisal powers of a reader becomes feasible; because of its separate identity, it can be arranged and rearranged in relation to other entries, thus maximizing the possibility of multiple classification of entries and variety of display patterns. And because it is machine-readable, all the manipulative powers of the computer can be used to optimize its use by scholars.

But the disadvantages of the STC entry can also be overcome by use of collocation files. Selection of those entries that require microscopic bibliographical control is perhaps the most important aspect of computer manipulation of the records (apart from classification and arrangement at the macroscopic level). Such selected entries will have minimal microscopic differentiations incorporated into the main file entry in the "points" field but an infinite extension of bibliographical notes can be incorporated into a collocation file. Perhaps the minimal amount of data of this kind would be a cross-reference to existing bibliographical catalogs, and a battery of separate collocation files for the most important of tools for microscopic bibliographical control might be the best approach to providing such linkages. Not only the HPB entry, but also the NUC entry would become the index term in such matching files, and the speed of searching and economy of display made possible by the computer would quickly bring the scholar into the interlocking bibliographical infrastructure of a gradually-evolving system for universal bibliographical control.

It is, therefore, important that a common reference system be adopted for developing those links. And the best one at hand is that of NUC: Pre-1956 Imprints. But can one imagine the immense amount of labor that would be

required to index in conventional ways the more than 600 volumes to which this work will ultimately extend? Even a chronological index of all the entries (which the tenor of my five lectures has perhaps urged to be the highest priority for scholars in the humanities) would be a lengthy task. But if it could be undertaken as a side-product of working on a yet more important priority - the improvement of coverage of items within the bibliographical universe of NUC - it might be worth undertaking. The development of an address system for the HPB/NUC Index does in fact accomplish this, for the 8-character address used in NUC is supplemented by a 4-character appendage which enables the last 3 figures of the date to be used technically as a checking device in the address, but also as a machine-readable datum for a chronological index to be created by computer.

The use of the four extra characters also permits the addition of new entries, cross-referencing of both new and old, and correction of misfiling in the huge printed work. Additions are indicated by an asterisk in the first of the four columns and a filing number in the last two. Existing cross-references are signalled by "00" in the last 2 columns, and new ones by the normal addition conventions.

It would not be possible to include additions in the automatic chronological index made from the NUC addresses alone, but as every addition must have an equivalent HPB entry (signalled in the matching file component of the HPB/NUC Index by an "=" sign in the 13th column), the date field of the HPB entry serves as an economical substitute.

An index of printers and publishers in NUC (or, more strictly, of names in the imprints) would not only be an enormous task for conventional indexing; it would be a most unsatisfactory tool for the scholar because the treatment of imprint data in traditional library cataloging is probably the weakest part of the data given in NUC. An index of printers and publishers for additions to NUC recorded in the HPB/NUC Index can be provided automatically. If scholars wished such an index to be extended to parts of NUC, then the entries in those parts should be given highest priority for HPB cataloging.

A subject index for the whole of NUC would be prohibitively expensive and somewhat arbitrary in coverage if undertaken by conventional manual procedures. A KWIC index (as less expensive but also less useful) could only be cost-effective if the data to be indexed were first put into machine-readable form - also a prohibitively expensive undertaking rejected by the RECON project investigators. The selective HPB cataloging of parts of the NUC listings might be undertaken as a side-product of inspection of particular items for limited subject classification.

Those items that fitted the subject or subjects chosen would be selected first by using the appropriate subject bibliographies already in existence, and then by matching them against NUC (the results of the matching would be incorporated into a collocation file for the subject, and into the HPB/NUC Index for the NUC/HPB matching element). Once the items were in the HPB/NUC Index, other forms of classification (chronological, printers and publishers, etc.) could be done by machine. The collocation file for the subject could, where it seemed useful, be in the form of a matching file where the NUC address was linked with the address system of a number of